

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 30, 2004, 15:22:40 ; Search time 59 Seconds
(without alignments)
1690.496 Million cell updates/sec

Title: US-09-466-778b-11

Perfect score: 1865
Sequence: 1 MTGPKHKCECKSHYVGDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003Bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1849	99.1	353	3 AAY93913	Aay93913 A human h
2	1772	95.0	1394	5 AAM47684	Aam47684 Human Hya
3	1772	95.0	1416	6 ABG72499	Abg72499 Human Hya
4	1772	95.0	1653	6 ABG72514	Abg72514 Human 190
5	1769	94.9	897	5 ABJ10588	Abj10588 Human nov
6	1554	83.3	2675	5 ABJ10586	Abj10586 Human nov
7	1514.5	81.2	1431	5 AAM47675	Aam47675 Rat Hyalu
8	1514.5	81.2	1431	6 ABG72498	Abg72498 Rat 175kd
9	1215.5	65.2	330	3 AAB42164	Aab42164 Human ORF
10	847	45.4	2420	5 ABJ10587	Abj10587 Human nov
11	843	45.2	315	4 AAB83358	Aab83358 NOV2 prot
12	661.5	35.5	1082	5 ABB90349	Abb90349 Human pol
13	658.5	35.3	2570	6 ABR82200	Abbr82200 Human CLE
14	656.5	35.2	457	3 AAY93911	Aay93911 A human h
15	656.5	35.2	2157	3 AAY93910	Aay93910 A human h
16	645.5	34.6	457	7 ADB47827	Adb47827 Novel hum
17	642	34.4	482	7 ADB65640	Adb65640 Human pro
18	581	31.2	510	7 ADB64426	Adb64426 Human pro
19	535	28.7	669	4 AAB83364	Aab83364 NOV6 prot
20	535	28.7	669	4 AAB83362	Aab83362 NOV6 prot
21	531	28.5	244	4 AAB83359	Aab83359 NOV3 prot
22	505	27.1	93	6 ABG72500	Abg72500 Human 190
23	262	14.0	277	2 AAR26044	Aar26044 Tumour ne
24	262	14.0	277	2 AAW13654	Aaw13654 Tumour ne
25	262	14.0	277	2 AAW84087	Aaw84087 Tumour ne

26	262	14.0	277	4	AAE02361	Human tum
27	262	14.0	277	5	ABG70870	Human tum
28	262	14.0	277	6	ABR58556	Human can
29	262	14.0	277	6	ABU56653	Lung can
30	262	14.0	277	7	ADE25781	Human pro
31	244	13.1	277	6	ABR58557	Human can
32	244	13.1	277	6	ABU56654	Lung can
33	213	11.4	106	4	AAU00893	Human can
34	189.5	10.2	289	3	AAAY93912	A human h
35	187.5	10.1	360	2	AAAY13381	Amino aci
36	187.5	10.1	360	3	ADC78533	Human PRO
37	187.5	10.1	360	4	AAAB80249	Human PRO
38	187.5	10.1	360	4	AAU29037	Human PRO
39	187.5	10.1	360	4	AAAM38965	Human pol
40	187.5	10.1	360	6	ABU58413	Human PRO
41	187.5	10.1	360	6	ABU71627	Human PRO
42	187.5	10.1	360	6	ABU87961	Novel hum
43	187.5	10.1	360	6	ABU84276	Human sec
44	187.5	10.1	360	6	ABR66150	Human sec
45	187.5	10.1	360	6	ABR65540	Human sec

ALIGNMENTS

RESULT 1

AAAY93913
ID AAY93913 standard; protein; 353 AA.

XX AAY93913;

XX AC

XX 03-OCT-2000 (first entry)

XX DE A human hyaluronan-binding protein, designated BM-HABP.

XX KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;

XX KW proliferative condition; metastasis; inflammation; ischemia;

XX KW host defence dysfunction; immune surveillance dysfunction; arthritis;

XX KW multiple sclerosis; autoimmunity; immune dysfunction; allergy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 94 /note= "unspecified amino acid encoded by NNC"

FT Domain 121..215

FT /note= "HA binding domain"

FT Misc-difference 157

FT /note= "unspecified amino acid encoded by TNC"

FT Misc-difference 303

FT /note= "unspecified amino acid encoded by YCT"

FT Misc-difference 314

FT /note= "unspecified amino acid encoded by NCC"

FT Misc-difference 320

FT /note= "unspecified amino acid encoded by GNA"

FT Misc-difference 324

FT /note= "unspecified amino acid encoded by TNT"

FT Misc-difference 325

FT /note= "unspecified amino acid encoded by GNC"

FT Misc-difference 351

FT /note= "unspecified amino acid encoded by CAN"

XX WO2000039166-A1.

XX 06-JUL-2000.

XX 20-DEC-1999; 99WO-US030462.

XX 23-DEC-1998; 98US-0113871P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (AMNA-) AMERICAN NAT RED CROSS.

XX

PI Hastings GA, Liau G, Tsiifrina E;
 XX WPI; 2000-452376/39.
 DR N-PSDB; AAA57365.
 XX
 PT New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-
 PT HABP and BM-HABP, useful for treating proliferative conditions,
 PT metastasis, inflammation, ischemia, arthritis and multiple sclerosis.
 XX
 PS Claim 11; Fig 4A-B; 457pp; English.
 XX
 CC The present sequence represents a hyaluronan-binding protein. The
 CC specification describes four hyaluronan-binding protein, known as WF-
 CC HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for
 CC treating diseases such as proliferative conditions, metastasis,
 CC inflammation, ischemia, host defence dysfunction, immune surveillance
 CC dysfunction, arthritis, multiple sclerosis, autoimmunity, immune
 CC dysfunction and allergy
 XX
 SQ Sequence 353 AA;
 Query Match 99.1%; Score 1849; DB 3; Length 353;
 Best Local Similarity 100.0%; Pred. No. 2.3e-184;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 60
 DB |||||
 QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 60
 DB |||||
 QY 61 RSPGQYKLTDFKAREACANEATWATYNOLSYXOKAYVHLCAGWLETGRVAYPTAFAS 120
 DB |||||
 QY 61 RSPGQYKLTDFKAREACANEATWATYNOLSYXOKAYVHLCAGWLETGRVAYPTAFAS 120
 DB |||||
 QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGVDGDFSYSGNLLQVLMGSP 180
 DB |||||
 QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGVDGDFSYSGNLLQVLMGSP 180
 DB |||||
 QY 181 SLTNFELTEVLAYSNSARGRAFLEHLDLSIRGTLPVQNSGLGENETLSGRDIEHHLAN 240
 DB |||||
 QY 181 SLTNFELTEVLAYSNSARGRAFLEHLDLSIRGTLPVQNSGLGENETLSGRDIEHHLAN 240
 DB |||||
 QY 241 VSMFFNDLVNGTTLQTRIGSKLLITDRQDPLHPTETRCVDRDTEWDICASNGITHVI 300
 DB |||||
 QY 241 VSMFFNDLVNGTTLQTRIGSKLLITDRQDPLHPTETRCVDRDTEWDICASNGITHVI 300
 DB |||||
 QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIILVTGAVALAAYSFRINRKTIGFXHF 353
 DB |||||
 QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIILVTGAVALAAYSFRINRKTIGFXHF 353
 DB |||||
 RESULT 2
 ID AAM47684
 AC AAM47684 standard; protein; 1394 AA.
 XX AAM47684;
 XX
 DT 22-FEB-2002 (first entry)
 DE Human Hyalurononic Acid Receptor for Endocytosis, HARE.
 XX
 KW HARE; Hyalurononic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
 KW chondroitin sulphate; extracellular matrix; cartilage; skin;
 KW vitreous humour; endocytic receptor; glycosaminoglycan; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200181544-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX 25-APR-2001; 2001WO-US013403.
 XX
 XX 25-APR-2000; 2000US-0199538P.
 XX

PR 02-NOV-2000; 2000US-0245320P.
 XX (WEIG/) WEIGEL P A.
 PA (ZHOU/) ZHOU B.
 PA (WEIG/) WEIGEL J A.
 XX Weigel PA, Zhou B, Weigel JA;
 PI WPI; 2002-049271/06.
 XX
 PT New mammalian hyalurononic acid receptor for endocytosis, useful e.g. for
 PT identifying agents that inhibit binding to hyalurononic acid, and related
 PT nucleic acid.
 XX
 PS Claim 20; Fig 33; 263pp; English.
 XX
 CC The present invention relates to sequences for rat and human HARE
 CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
 CC and AAM47684). HARE can bind specifically to at least one of hyalurononic
 CC acid (HA), also known as hyaluronan), chondroitin (C) or chondroitin
 CC sulphate (CS). HA is an extracellular matrix component of all tissues; in
 CC particular cartilage, skin and vitreous humour. HARE is the endocytic
 CC receptor responsible for removing HA and other glycosaminoglycans from
 CC the circulation. The present sequence is human HARE
 XX
 SQ Sequence 1394 AA;
 Query Match 95.0%; Score 1772; DB 5; Length 1394;
 Best Local Similarity 94.6%; Pred. No. 2.1e-175;
 Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;
 QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 60
 DB |||||
 QY 986 MTGPGKHCKESHVYVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 1045
 DB |||||
 QY 61 RSPGQYKLTDFKAREACANEATWATYNOLSYXOKAYVHLCAGWLETGRVAYPTAFAS 120
 DB |||||
 QY 1046 RSPGQYKLTDFKAREACANEATWATYNOLSYXOKAYVHLCAGWLETGRVAYPTAFAS 1105
 DB |||||
 QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGVDGDFSYSGNLLQVLMGSP 180
 DB |||||
 QY 1106 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGVDGDFSYSGNLLQVLMGSP 1165
 DB |||||
 QY 181 SLTNFELTEVLAYSNSARGRAFLEHLDLSIRGTLPVQNSGLGENETLSGRDIEHHLAN 240
 DB |||||
 QY 1166 SLTNFELTEVLAYSNSARGRAFLEHLDLSIRGTLPVQNSGLGENETLSGRDIEHHLAN 1225
 DB |||||
 QY 241 VSMFFNDLVNGTTLQTRIGSKLLITDRQDPLHPTETRCVDRDTEWDICASNGITHVI 300
 DB |||||
 QY 1226 VSMFFNDLVNGTTLQTRIGSKLLITASQDPLQPTETRFVDRALQWDIFASNGIITHVI 1285
 DB |||||
 QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIILVTGAVALAAYSFRINRKTIGFXHF 353
 DB |||||
 QY 1286 SRPLKAPPAPVTLTHTGLGAGIFFAIIIVTGAVALAAYSFRINRKTIGFQHF 1338
 DB |||||
 RESULT 3
 ID ABG72499
 AC ABG72499 standard; protein; 1416 AA.
 XX ABG72499;
 XX
 DT 19-FEB-2003 (first entry)
 DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #1.
 XX
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200286093-A2.

```

XX PD 31-OCT-2002.
XX AC
XX AC ABG72514 standard; protein; 1653 AA.
XX AC ABG72514;
XX DT 19-FEB-2003 (first entry)
XX DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #2.
XX KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
XX KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
XX KW tumour; gene therapy; human.
XX OS Homo sapiens.
XX PN WO200286093-A2.
XX PD 31-OCT-2002.
XX PF 25-APR-2002; 2002WO-US013209.
XX PR 25-APR-2001; 2001US-00842930.
XX PR 25-APR-2001; 2001US-0286460P.
XX XX (WEIG/) WEIGEL P H.
XX PA (WEIG/) WEIGEL J A.
XX XX Weigel PH, Weigel JA;
XX DR WPI; 2003-093126/08.
XX DR N-PSDB; ABX13823.
XX PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
XX PT expressing functional active hyaluronan receptor for endocytosis of HARE,
XX PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
XX PT epitope.
XX XX Example; Fig. 9A; 167pp; English.
XX CC The invention describes a method of targeting a compound to a cell or
XX CC tissue of an individual expressing a functionally active hyaluronan (HA)
XX CC receptor for endocytosis (HARE) or a cell that does not express
XX CC functionally active HARE. The method involves using HA molecule, a
XX CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
XX CC a monoclonal antibody raised against a HA-binding domain of HARE. The
XX CC method is useful for targeting a compound, preferably a chemotherapeutic
XX CC agent or a radioisotope to cell of an individual, especially a human,
XX CC expressing HARE on its surface (e.g. gene therapy). Also described is a
XX CC method useful for preventing interaction between a cell having at least
XX CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
XX CC its surface. This second method is useful for preventing metastasis by
XX CC non-tumour cells expressing HARE on its surface. The invention also
XX CC describes a method useful for detecting the presence of HA, CD and CDS in
XX CC a biological fluid. This sequence encodes the human 190kDa Hyaluronan
XX CC receptor for endocytosis (HARE)
XX SQ Sequence 1416 AA;
Query Match 95.0%; Score 1772; DB 6; Length 1416;
Best Local Similarity 94.6%; Pred. No. 2.1e-175;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;
QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 1008 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 1067
QY 61 RSPFGQYKLTFDKAREACANEATWATYNQLSYQAKYHLCSAGWLETGRVAYTAFAS 120
Db 1068 RSPFGQYKLTFDKAREACANEATWATYNQLSYQAKYHLCSAGWLETGRVAYTAFAS 1127
QY 121 QNCGSGVGVINDYGPBNKSEMDVFCYRMKDVNCTXKVGVDGFSYGNLLQVLMSPF 180
Db 1128 QNCGSGVGVINDYGPBNKSEMDVFCYRMKDVNCTXKVGVDGFSYGNLLQVLMSPF 1187
QY 181 SLTNFLTEVLAYSNSSGRAFLHLELTLISIRGLTFVPQNSGLGENETLSGRDIEHHLAN 240
Db 1188 SLTNFLTEVLAYSNSSGRAFLHLELTLISIRGLTFVPQNSGLGENETLSGRDIEHHLAN 1247
QY 241 VSMFFNDLVNGTTLQRLGSKLLITDRQDPLHPTETRCVGRDRLTLEWDICASNGITHVI 300
Db 1248 VSMFFNDLVNGTTLQRLGSKLLITASQDPLQPTETRFVDRAILQWDIFASNGIIHVI 1307
QY 301 SRXLKAPPAPVTLXHTGLGIFXIIILVTGVALAAYSVERINRXTIGFQHF 353
Db 1308 SRPLKAPPAPVTLXHTGLGIFXIIILVTGVALAAYSVERINRRTIGFQHF 1360

```

RESULT 4

```

ABG72514
ID ABG72514 standard; protein; 1653 AA.
XX AC
XX AC ABG72514;
XX DT 19-FEB-2003 (first entry)
XX DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #2.
XX KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
XX KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
XX KW tumour; gene therapy; human.
XX OS Homo sapiens.
XX PN WO200286093-A2.
XX PD 31-OCT-2002.
XX PF 25-APR-2002; 2002WO-US013209.
XX PR 25-APR-2001; 2001US-00842930.
XX PR 25-APR-2001; 2001US-0286460P.
XX XX (WEIG/) WEIGEL P H.
XX PA (WEIG/) WEIGEL J A.
XX XX Weigel PH, Weigel JA;
XX DR WPI; 2003-093126/08.
XX DR N-PSDB; ABX13823.
XX PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
XX PT expressing functional active hyaluronan receptor for endocytosis of HARE,
XX PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
XX PT epitope.
XX XX Example; Fig. 9B; 167pp; English.
XX CC The invention describes a method of targeting a compound to a cell or
XX CC tissue of an individual expressing a functionally active hyaluronan (HA)
XX CC receptor for endocytosis (HARE) or a cell that does not express
XX CC functionally active HARE. The method involves using HA molecule, a
XX CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
XX CC a monoclonal antibody raised against a HA-binding domain of HARE. The
XX CC method is useful for targeting a compound, preferably a chemotherapeutic
XX CC agent or a radioisotope to cell of an individual, especially a human,
XX CC expressing HARE on its surface (e.g. gene therapy). Also described is a
XX CC method useful for preventing interaction between a cell having at least
XX CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
XX CC its surface. This second method is useful for preventing metastasis by
XX CC non-tumour cells expressing HARE on its surface. The invention also
XX CC describes a method useful for detecting the presence of HA, CD and CDS in
XX CC a biological fluid. This is the amino acid sequence of a longer version
XX CC of the human 190kDa Hyaluronan receptor for endocytosis (HARE) shown in
XX CC ABG72439 (Encoded by ABX13822)
XX SQ Sequence 1653 AA;
Query Match 95.0%; Score 1772; DB 6; Length 1653;
Best Local Similarity 94.6%; Pred. No. 2.7e-175;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;
QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 1245 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 1304
QY 61 RSPFGQYKLTFDKAREACANEATWATYNQLSYQAKYHLCSAGWLETGRVAYTAFAS 120
Db 1305 RSPFGQYKLTFDKAREACANEATWATYNQLSYQAKYHLCSAGWLETGRVAYTAFAS 1364
QY 121 QNCGSGVGVINDYGPBNKSEMDVFCYRMKDVNCTXKVGVDGFSYGNLLQVLMSPF 180

```

Db 1365 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFCGNNLLQVLMSPF 1424
 Qy 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 Db 1425 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1484
 Qy 241 VSMFFYNLNVNGTTLTQRLGSKLLITDRQDPLHPTETRCVDCGRDLEWDICASNGITHVI 300
 Db 1495 VSMFFYNLNVNGTTLTQRLGSKLLITASQDPLQPTETRFVDGRLAQWDIFASNGLIHVI 1544
 Qy 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFYFRINRKTIGFXHF 353
 Db 1545 SRPLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSFYFRINRKTIGFQHF 1597

RESULT 5

ABJ10588
 ID ABJ10588 standard; protein; 897 AA.

AC ABJ10588;

XX 28-NOV-2002 (first entry)

XX Human novel protein NOV1c SEQ ID NO: 211.

XX Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
 KW virucide; tranquilizer; anticoagulant; osteopathic; analgesic;
 KW antiparkinsonian; dermatological; antifertility; cerebroprotective;
 KW antiaddictive.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 221
 FT /note= "optionally Cys depending on polymorphism present
 in coding sequence"
 FT Misc-difference 325
 FT /note= "optionally Leu depending on polymorphism present
 in coding sequence"
 FT Misc-difference 416
 FT /note= "optionally Tyr depending on polymorphism present
 in coding sequence"
 FT Misc-difference 832
 FT /note= "optionally Gly depending on polymorphism present
 in coding sequence"

XX WO200259315-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US0500076.

XX 19-DEC-2000; 2000US-0256619P.

XX 19-JAN-2001; 2001US-0262959P.

XX 28-FEB-2001; 2001US-0272408P.

XX 20-APR-2001; 2001US-0285189P.

XX 26-JUL-2001; 2001US-0308039P.

XX 09-AUG-2001; 2001US-0311266P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;

PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;

PI Taupier RJ, Kekuda R, Smithson G, Zerkhusen BD, Liu X, Colman SD;

PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;

PI Rothenberg M;

XX WPI; 2002-666903/71.

DR N-PSDB; AB08489.

XX

XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
 PT or Alzheimer's disease.
 XX Claim 54; Page 25; 363pp; English.

XX The present invention provides the protein and coding sequences of
 CC several novel human proteins, designated NOVX. These can be used in the
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberos sclerosis,
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 CC palsy, epilepsy, Lesh-Nyhan syndrome, multiple sclerosis, ataxia,
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
 CC graft-versus-host disease. The present sequence is a protein of the
 CC invention
 XX

SQ Sequence 897 AA;

Query Match 94.9%; Score 1769; DB 5; Length 897;

Best Local Similarity 94.3%; Pred. No. 2.2e-175;

Matches 333; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MTGPKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITVGVPHL 60

Db 489 MTGPKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITVGVPHL 548

Qy 61 RSPLQYKLTDFDKAREACANEAATWATYNQLSYOKAKYHLCSAGWLETGRVAYPTAFAS 120

Db 549 RSPLQYKLTDFDKAREACANEAATWATYNQLSYOKAKYHLCSAGWLETGRVAYPTAFAS 608

Qy 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFCGNNLLQVLMSPF 180

Db 609 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFCGNNLLQVLMSPF 568

Qy 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240

Db 669 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 728

Qy 241 VSMFFYNLNVNGTTLTQRLGSKLLITDRQDPLHPTETRCVDCGRDLEWDICASNGITHVI 300

Db 729 VSMFFYNLNVNGTTLTQRLGSKLLITASQDPLQPTETRFVDGRLAQWDIFASNGLIHVI 788

Qy 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFYFRINRKTIGFXHF 353

Db 789 SRPLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSFYFRINRKTIGFQHF 841

RESULT 6

ABJ10586

ID ABJ10586 standard; protein; 2675 AA.

XX ABJ10586;

XX 28-NOV-2002 (first entry)

XX Human novel protein NOV1a SEQ ID NO: 2.

XX Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
 KW virucide; tranquilizer; anticoagulant; osteopathic; analgesic;
 KW antiparkinsonian; dermatological; antifertility; cerebroprotective;
 KW antiaddictive.

XX Homo sapiens.

EN WO200259315-A2.
 XX 01-AUG-2002.
 XX 19-DEC-2001; 2001WO-US050076.
 XX 19-DEC-2000; 2000US-0256619P.
 XX 19-JAN-2001; 2001US-0262959P.
 XX 28-FEB-2001; 2001US-0272408P.
 XX 20-APR-2001; 2001US-0285189P.
 XX 26-JUL-2001; 2001US-0308039P.
 XX 09-AUG-2001; 2001US-0311266P.
 XX (CURA-) CURAGEN CORP.
 XX Shinkets RA, Patturajan M, Vernet CAM, Casman SU, Malyankar U;
 XX Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
 XX Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
 XX Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
 XX Rothenberg M;
 XX WPI; 2002-666903/71.
 XX N-PSDB; ABT08487.
 XX New isolated NOVX polypeptides and polynucleotides, useful for
 XX preventing, diagnosing or treating NOVX-associated disorders e.g.
 XX diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
 XX or Alzheimer's disease.
 XX Claim 1; Page 20; 363pp; English.
 XX The present invention provides the protein and coding sequences of
 XX several novel human proteins, designated NOVX. These can be used in the
 XX treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
 XX Lindau syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
 XX hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 XX palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
 XX telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
 XX obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
 XX infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
 XX diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
 XX immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
 XX graft-versus-host disease. The present sequence is a protein of the
 XX invention
 XX SQ Sequence 2675 AA;
 Query Match 83.3%; Score 1554; DB 5; Length 2675;
 Best Local Similarity 83.9%; Pred. No. 3.6e-152;
 Matches 296; Conservative 5; Mismatches 16; Indels 36; Gaps 1;
 QY 1 MTGPGKHCKECKSHYVGDGLNCPPEQLPDRCLQDNGQCHADAKCVDLHPQDITVGVFHL 60
 Db 2303 MTGPGKHCKECKSHYVGDGLNCPPEQLPDRCLQDNGQCHADAKCVDLHPQDITVGVFHL 2362
 QY 61 RSLPGYKLTDFKAREACANEATMATYNOLSYOKAKVHLCAGWLETGRVAYPTAFAS 120
 Db 2363 RSLPGYKLTDFKAREACANEATMATYNOLSYOKAKVHLCAGWLETGRVAYPTAFAS 2422
 QY 121 QNCGSGVGVGVVDYGPGRNPKSEMDVFCYRMKDVNCTKXKVGVGDGFSYGNLLQVLSFP 180
 Db 2423 QNCGSGVGVGVVDYGPGRNPKSEMDVFCYRMK- - - - - 2453
 QY 181 SLTNFLTEVLAYSNSARGRAFLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 Db 2454 - - - - - EVLAYSNSARGRAFLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 2506
 QY 241 VSMFFYNDLVNGTTLQTRLSKLLITDROPLHPTETRCVDGRTLEWDICASNGITHVI 300
 Db 2507 VSMFFYNDLVNGTTLQTRLSKLLITASQDPLQVQSRFVDGRAILQWDIPASNGIITHVI 2566
 QY 301 SRXLKAPPAPVTLTHTGLGIGIFXKXIIIVTGAVALAAYSYPRIINRTIGYQHF 353

Db 2567 SRPLKAPPAPVTLTHTGLGIGIFXKXIIIVTGAVALAAYSYPRIINRTIGYQHF 2619
 RESULT 7
 AAM47675
 ID AAM47675 standard; protein; 1431 AA.
 XX AAM47675;
 XX 22-FEB-2002 (first entry)
 XX Rat Hyaluronic Acid Receptor for Endocytosis, HARE.
 XX HARE: Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
 XX chondroitin sulphate; extracellular matrix; cartilage; skin;
 XX vitreous humour; endocytic receptor; glycosaminoglycan; rat.
 XX Rattus norvegicus.
 XX WO200181544-A2.
 XX 01-NOV-2001.
 XX 25-APR-2001; 2001WO-US013403.
 XX 25-APR-2000; 2000US-0195538P.
 XX 02-NOV-2000; 2000US-0245320P.
 XX (WEIG/) WEIGEL P A.
 XX (ZHOU/) ZHOU B.
 XX (WEIG/) WEIGEL J A.
 XX Weigel PA, Zhou B, Weigel JA;
 XX WPI; 2002-049271/06.
 XX N-PSDB; ABA04648.
 XX New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
 XX identifying agents that inhibit binding to hyaluronic acid, and related
 XX nucleic acid.
 XX Claim 20; Fig 21; 263pp; English.
 XX The present invention relates to sequences for rat and human HARE
 XX (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
 XX and AAM47684). HARE can bind specifically to at least one of hyaluronic
 XX acid (HA), also known as hyaluronan), chondroitin (C) or chondroitin
 XX sulphate (CS). HA is an extracellular matrix component of all tissues, in
 XX particular cartilage, skin and vitreous humour. HARE is the endocytic
 XX receptor responsible for removing HA and other glycosaminoglycans from
 XX the circulation. The present sequence is rat HARE
 XX SQ Sequence 1431 AA;
 Query Match 81.2%; Score 1514.5; DB 5; Length 1431;
 Best Local Similarity 79.3%; Pred. No. 1.9e-148;
 Matches 280; Conservative 29; Mismatches 43; Indels 1; Gaps 1;
 QY 1 MTGPGKHCKECKSHYVGDGLNCPPEQLPDRCLQDNGQCHADAKCVDLHPQDITVGVFHL 60
 Db 1008 MTGPGKHCKECKSHYVGDGLNCPPEQLPDRCLQDNGQCHADAKCVDLHPQDITVGVFHL 1067
 QY 61 RSLPGYKLTDFKAREACANEATMATYNOLSYOKAKVHLCAGWLETGRVAYPTAFAS 120
 Db 1068 RSLPGYKLTDFKAREACANEATMATYNOLSYOKAKVHLCAGWLETGRVAYPTAFAS 1127
 QY 121 QNCGSGVGVGVVDYGPGRNPKSEMDVFCYRMKDVNCTKXKVGVGDGFSYGNLLQVLSFP 180
 Db 1128 QKCGANVGVGVVDYGPGRNPKSEMDVFCYRMKDVNCTKXKVGVGDGFSYGNLLQVLSFP 1187
 QY 181 SLTNFLTEVLAYSNSARGRAFLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 Db 1188 SLTNFLTEVLAYSNSARGRAFLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1247

QY 241 VSMFFYNDLVNGTTLQTRIGSKLLITDRQDPLHPETRCVDCGRDLEWIDICASNIGITHVI 300
 DB 1248 VNVSYNDLVNGTFLRTMLGSQLITFSQDLH-QETRFVDCRSILOWDIIAANGILHII 1306
 QY 301 SXRLKAPPAPVTLXHTGLGCGIFXXIILVTGAVALAAASYFRINRKTIGFXHF 353
 DB 1307 SEPLRAPPTAATAAHSGLGTGIFCAVLVTGAIALAAASYFRILKQRTTGQRF 1359

RESULT 8

ABG72498
 ID ABG72498 standard; protein; 1431 AA.

XX
 AC ABG72498;

XX
 DT 19-FEB-2003 (first entry)

XX
 DE Rat 175KDa Hyaluronan receptor for endocytosis (HARE).

XX
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; rat; receptor.

XX
 OS Rattus sp.

XX
 FN WO200286093-A2.

XX
 PD 31-OCT-2002.

XX
 PF 25-APR-2002; 2002WO-US013209.

XX
 PR 25-APR-2001; 2001US-00842930.

XX
 PR 25-APR-2001; 2001US-0286468P.

XX
 PA (WEIG/) WEIGEL P H.

XX
 PA (WEIG/) WEIGEL J A.

XX
 PI Weigel PH, Weigel JA;

XX
 DR WPI; 2003-093126/08.

XX
 DR N-PSDB; ABX13821.

XX
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject

PT
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.

XX
 PS Example; Fig 2; 167pp; English.

XX
 CC The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This is the amino acid sequence of the rat 175KDa
 CC Hyaluronan receptor for endocytosis (HARE)

XX
 SQ Sequence 1431 AA;

Query Match 81.2%; Score 1514.5; DB 6; Length 1431;

Best Local Similarity 79.3%; Pred. No. 1.9e-148;

Matches 280; Conservative 29; Mismatches 43; Indels 1; Gaps 1;

QY 1 MTGPKHKCECKSHYVGVGDLNCEPEQLPIDRCLQDNGOCHADAKVLDLHFODTTVGVFHL 60
 DB 1008 MTGPKHKCECKSHYVGVGDLNCEPEQLPIDRCLQDNGOCHADAKVLDLHFODTTVGVFHL 1067
 QY 61 RSPGLGOYKLTDFDKAREACANEAAATMATYNQLSYOKAKYHLCAGWLETGRVAYPTAFAS 120
 DB 1068 RSPGLGOYKLTDFDKAREACANEAAATMATYNQLSYOKAKYHLCAGWLETGRVAYPTAFAS 1127
 QY 121 QNCGSVVGIIVDYGPRPNKSEMDVFCYRMKDVCNCKXGVYVGVGDFSYGNLLQVLMSPF 180
 DB 1128 QNCGSVVGIIVDYGPRPNKSEMDVFCYRMKDVCNCKXGVYVGVGDFSYGNLLQVLMSPF 1187
 QY 181 SLTNFLTEVLAYSNSAARGRAFLEHLTDLSTRTGLFVPQNSGLGENETLSGEDIHHLAN 240
 DB 1188 SLTNFLTEVLAYSNSAARGRAFLEHLTDLSTRTGLFVPQNSGLGENETLSGEDIHHLAN 1247
 QY 241 VSMFFYNDLVNGTTLQTRIGSKLLITDRQDPLHPETRCVDCGRDLEWIDICASNIGITHVI 300
 DB 1248 VNVSYNDLVNGTFLRTMLGSQLITFSQDLH-QETRFVDCRSILOWDIIAANGILHII 1306
 QY 301 SXRLKAPPAPVTLXHTGLGCGIFXXIILVTGAVALAAASYFRINRKTIGFXHF 353
 DB 1307 SEPLRAPPTAATAAHSGLGTGIFCAVLVTGAIALAAASYFRILKQRTTGQRF 1359

RESULT 9

AAB42164

ID AAB42164 standard; protein; 330 AA.

XX
 AC AAB42164;

XX
 DT 08-FEB-2001 (first entry)

XX
 DE Human ORFX ORF1928 polypeptide sequence SEQ ID NO:3856.

XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineoplastic; antitumor;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX
 OS Homo sapiens.

XX
 WO2000058473-A2.

XX
 PD 05-OCT-2000.

XX
 PF 31-MAR-2000; 2000WO-US008621.

XX
 PR 31-MAR-1999; 99US-0127607P.

XX
 PR 02-APR-1999; 99US-0127636P.

XX
 PR 05-APR-1999; 99US-0127728P.

XX
 PR 30-MAR-2000; 2000US-00540763.

XX
 PA (CURA-) CURAGEN CORP.

XX
 PI Shimkets RA, Leach M;

XX
 DR WPI; 2000-602362/57.

XX
 DR N-PSDB; AAC76373.

XX
 PT Novel nucleic acids and peptides derived from open reading frame X,

PT useful for treating e.g. cancers, proliferative disorders,

RESULT 10	
ABJ10587	
ID	ABJ10587 standard; protein; 2420 AA.
XX	
XX	AC
XX	ABJ10587;
XX	
XX	
DT	28-NOV-2002 (first entry)
XX	
XX	
DE	Human novel protein NOV1b SEQ ID NO: 4.
XX	
KW	Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW	antiarteriosclerotic; antidiabetic; antilasthmatic; antiinflammatory;
KW	haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW	antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW	varicude; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW	antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW	antiaddictive.
XX	
XX	
XX	Homc sapiens.
OS	
XX	
XX	
PN	WO200259315-A2.
PN	

AA883358;

26-MAR-2002 (first entry)

NOV2 protein sequence.

NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory; cardiovascular; casein kinase II phosphorylation site; contraception; serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation; epidermal growth factor; cell development; apoptosis; cell adhesion; growth migration; cell structure; motility; cancer; immune disorder; inflammatory disorder; cellular adhesion disorder; long-QT syndrome; cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome; therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10; NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.

Unidentified.

WO200136638-A2.

25-MAY-2001.

17-NOV-2000; 2000WO-US031543.

19-NOV-1999; 93US-0166336P.

29-NOV-1999; 93US-0167785P.

08-MAR-2000; 2000US-0187844P.

16-NOV-2000; 2000US-00715417.

(CURA-) CURAGEN CORP.

Shimkets RA, Lichenstein H, Vernet C, Fernandes E;

WPI; 2001-648134/74.

N-PSDB; AAF87113.

Novel human polypeptides and the nucleic acids that encode them useful for preventing, diagnosing and treating e.g. cancer, inflammation and immune disorders.

Claim 1; Page 10-13; 141pp; English.

This sequence is the NOV2 protein. The invention relates to the NOV1-NOV16 proteins, and their coding sequences. The proteins have Cytostatic; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular activities. The sequences may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate NOVX expression. They may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of protein by expressing inactive proteins or to supplement the patients own production of protein. They are used to produce NOVX proteins, by inserting the nucleic acid into a cell and culturing it to express the protein. The DNA may be used as DNA probes in assays to detect and quantitate the presence of similar DNAs in samples, and which patients may need restorative therapy. The NOVX protein may also be used as antigens in the production of antibodies (Abs) against NOVX and in assays to identify modulators of NOVX expression and activity. The anti-NOVX Abs and antagonist are used to down regulate expression and activity. The anti-NOVX Abs are used for detecting the presence of NOVX in samples. Disorders that may be prevented, diagnosed and/or treated vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and NOV13-16 have casein kinase II phosphorylation sites characteristic of serine/threonine kinases, and are used to treat kinase-related disorders (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception). NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF) like super family and are involved in, e.g. regulation of cell development, apoptosis, cell adhesion, growth migration, cell structure and motility and protein management, and are used to treat cancers, inflammatory disorders, immune disorders and cellular adhesion disorders. NOV6-10 are homologous to EGF-like fibrillin proteins and are used to treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT syndrome and marfan syndrome

Sequence 315 AA;

Query Match 45.2%; Score 843; DB 4; Length 315;
Best Local Similarity 84.7%; Pred. No. 3.8e-79;
Matches 155; Conservative 2; Mismatches 10; Indels 16; Gaps 1;

QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFDQTTWGVFHL 60
DB 145 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFDQTTWGVFHL 204
QY 61 RSPLGQYKLTDFDKARACANEAAATMATYNQLSYXQKAKYHLCSAGWLETGRVAYPTAPAS 120
DB 205 RSPLGQYKLTDFDKARACANEAAATMATYNQLSYXQKAKYHLCSAGWLETGRVAYPTAPAS 264
QY 121 QNCGSGVGVIVDYGPRNKSEMDVFCYRKNVNTYKXGVGVGDFSPYSYGNLLQVLMSPF 180
DB 265 QNCGSGVGVIVDYGPRNKSEMDVFCYRKM-----SAGLFOQLSSRP 308
QY 181 SLT 183
DB 309 CIS 311

RESULT 12

ABB90349

ID ABB90349 standard; protein; 1082 AA.

XX ABB90349;

AC ABB90349;

DT 24-MAY-2002 (first entry)

DE Human polypeptide SEQ ID NO 2725.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX WO200190304-A2.

PD 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US016450.

PR 19-MAY-2000; 2000US-0205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

DR N-PSDB; ABL90758.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders.

XX Claim 11; SEQ ID NO 2725; 2081pp + Sequence Listing; English.

PS The invention relates to novel genes (ABL89449-ABL90853) and proteins

CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in

CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and

CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,

CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune

CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic

CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)

CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing

PR XX PA PA PA XX PI

AA DR DR XX

XX PD PD PD

CC CC XX CC

U U U
U U U

U U U
U U U

U U U

85

Qy

D'b	Qy
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Dib

Qy

Qy

Db Qy

Qy

D**b**

Query Match 35.2%; Score 656.5; DB 3; Length 457;
Best Local Similarity 39.7%; Pred. No. 2.2e-59;
Matches 144; Conservative 50; Mismatches 148; Indels 21; Gaps 7;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 30, 2004, 15:37:50 ; Search time 45 Seconds
(without alignments)
2475.065 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPKHKCKECKSHYVDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1772	95.0	1192	4 Q9H7H7	Q9H7H7 homo sapien
2	1772	95.0	1416	4 Q86UR4	Q86UR4 homo sapien
3	1772	95.0	1736	4 Q8TES1	Q8TES1 homo sapien
4	1772	95.0	2551	4 Q8WQ08	Q8WQ08 homo sapien
5	1772	95.0	2551	4 Q8IUG9	Q8IUG9 homo sapien
6	1772	95.0	2551	4 Q7Z5N9	Q7Z5N9 homo sapien
7	1769	94.9	897	4 Q9NRY3	Q9NRY3 homo sapien
8	1514.5	81.2	1431	11 Q8CFM6	Q8CFM6 rattus norv
9	1487	79.7	894	11 Q8BM87	Q8BM87 mus musculu
10	1487	79.7	2559	11 Q8R4U0	Q8R4U0 mus musculu
11	1480	79.4	1069	4 Q9UF98	Q9UF98 homo sapien
12	664.5	35.6	1238	11 Q8K0K6	Q8K0K6 mus musculu
13	664.5	35.6	2571	11 Q8R4Y4	Q8R4Y4 mus musculu
14	661.5	35.5	2570	4 Q8IUH1	Q8IUH1 homo sapien
15	661.5	35.5	2589	4 Q93072	Q93072 homo sapien
16	652.5	35.0	2570	4 Q9NY15	Q9NY15 homo sapien

17	643	34.5	500	6 Q864U4	Q864U4 bos taurus
18	324.5	17.4	174	11 Q8VC09	Q8VC09 mus musculu
19	187.5	10.1	360	4 Q96S86	Q96S86 homo sapien
20	178.5	9.6	359	11 Q80XX3	Q80XX3 mus musculu
21	178.5	9.6	359	11 Q80WM5	Q80WM5 mus musculu
22	177	9.5	2109	13 P79787	P79787 gallus gall
23	173	9.3	494	6 Q9BGH3	Q9BGH3 sus scrofa
24	171.5	9.2	400	11 Q80WM4	Q80WM4 mus musculu
25	171.5	9.2	400	11 Q80XX2	Q80XX2 mus musculu
26	164.5	8.8	402	4 Q86UW8	Q86UW8 homo sapien
27	164.5	8.8	412	4 Q86PW2	Q86PW2 homo sapien
28	163.5	8.8	1290	13 Q9W6E1	Q9W6E1 gallus gall
29	162.5	8.7	359	13 Q7ZX17	Q7ZX17 xenopus lae
30	155	8.3	516	4 Q81WX2	Q81WX2 homo sapien
31	154.5	8.3	671	4 Q9HBK1	Q9HBK1 homo sapien
32	154.5	8.3	911	4 Q9HBK4	Q9HBK4 homo sapien
33	154.5	8.3	911	4 Q96FF7	Q96FF7 homo sapien
34	154.5	8.3	911	4 Q8TBB9	Q8TBB9 homo sapien
35	154.5	8.3	911	4 Q96GM7	Q96GM7 homo sapien
36	153	8.2	192	6 Q02817	Q02817 oryctolagus
37	150.5	8.1	354	4 Q86W61	Q86W61 homo sapien
38	150.5	8.1	883	11 Q80WT7	Q80WT7 mus musculu
39	150	8.0	1152	13 Q90WM2	Q90WM2 xenopus lae
40	144	7.7	210	11 Q8C9U7	Q8C9U7 mus musculu
41	144	7.7	368	11 Q8BS97	Q8BS97 mus musculu
42	128	6.9	297	5 Q7YUE5	Q7YUE5 herdmania c
43	126.5	6.8	103	6 Q9TTB3	Q9TTB3 sus scrofa
44	120.5	6.5	103	6 Q46380	Q46380 oryctolagus
45	120	6.4	2360	5 Q7YZP0	Q7YZP0 elmeria max

ALIGNMENTS

RESULT 1

Q9H7H7
ID Q9H7H7 PRELIMINARY; PRT; 1192 AA.
AC Q9H7H7;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE FLJ00112 protein (Fragment).
GN FLJ00112.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Splicein;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human spleen."
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK024503; BAB15793.1; -
DR HSSP; P98066; 1TSG.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 7.
DR Pfam; PF02469; Fasciclin; 2.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 9.
DR SMART; SM00554; FAS1; 3.
DR SMART; SM00445; LINK; 1.

```

DR PROSITE; PS00022; EGF_1; 4.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS02013; FAS1; 3.
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
KW EGF-like domain; Laminin EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1192 AA; 128738 MW; 59668ED4B83BD9C1 CRC64;

Query Match          95.0%; Score 1772; DB 4; Length 1192;
Best Local Similarity 94.6%; Pred. No. 4e-158;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFPHL 60
DB 784 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFPHL 843
QY 61 RSLPGQYKLTDPKAREACANEATMATYNQLSYAKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 844 RSLPGQYKLTDPKAREACANEATMATYNQLSYAKAKYHLCSAGWLETGRVAYPTAFAS 903
QY 121 QNCGGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGVDGFGFSYGNLLQVLMSPF 180
DB 904 QNCGGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVDGFGFSYGNLLQVLMSPF 963
QY 181 SLTNFLTEVLAYSNSARGRAFLHLDLTLIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
DB 964 SLTNFLTEVLAYSNSARGRAFLHLDLTLIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1023
QY 241 VSMFFYNDLVNGTTLQTRLSKLLITDRQDPLHPTETRCVDGRDITLWDICASNGITHVI 300
DB 1024 VSMFFYNDLVNGTTLQTRLSKLLITASQDPLQPTETRFVDRGAILQWIDIFASNGIITHVI 1083
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIILVTGAVALAAYSYFRINRKTIGFXHF 353
DB 1084 SRPLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSYFRINRRTTIGFQHF 1136

RESULT 2
Q86UR4 PRELIMINARY; PRT; 1416 AA.
AC Q86UR4;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hyaluronan receptor for endocytosis (Fragment).
GN HARE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22588541; PubMed=12626425;
RA Zhou B., McGary C.T., Weigel J.A., Saxena A., Weigel P.H.;
RT "Purification and molecular identification of the human hyaluronan
RT receptor for endocytosis."
RL Glycobiology 13:339-349(2003).
DR EMBL; AY227444; AAC39681.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR InterPro; IPR001673; S_mold_repeat.
DR Pfam; PF00008; EGF; 8.
DR Pfam; PF02469; Fasciclin; 3.
DR Pfam; PF00193; Xlink; 1.

PRINTS; PR00011; EGF_LAMININ.
ProDom; PD000918; Link; 1.
ProDom; PD06869; S_mold_repeat; 1.
SMART; SM00181; EGF; 11.
SMART; SM00179; EGF_CA; 4.
SMART; SM00554; FAS1; 4.
SMART; SM00445; LINK; 1.
PROSITE; PS00022; EGF_1; 4.
PROSITE; PS01186; EGF_2; 8.
PROSITE; PS02013; FAS1; 3.
PROSITE; PS01248; LAMININ TYPE EGF; 2.
KW Receptor.
FT NON_TER 1
FT CHAIN <1 1416
SQ SEQUENCE 1416 AA; 154089 MW; 85A216D38E3B10DE CRC64;

Query Match          95.0%; Score 1772; DB 4; Length 1416;
Best Local Similarity 94.6%; Pred. No. 5.1e-158;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFPHL 60
DB 1008 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFPHL 1067
QY 61 RSLPGQYKLTDPKAREACANEATMATYNQLSYAKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 1068 RSLPGQYKLTDPKAREACANEATMATYNQLSYAKAKYHLCSAGWLETGRVAYPTAFAS 1127
QY 121 QNCGGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGVDGFGFSYGNLLQVLMSPF 180
DB 1128 QNCGGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVDGFGFSYGNLLQVLMSPF 1187
QY 181 SLTNFLTEVLAYSNSARGRAFLHLDLTLIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
DB 1188 SLTNFLTEVLAYSNSARGRAFLHLDLTLIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1247
QY 241 VSMFFYNDLVNGTTLQTRLSKLLITDRQDPLHPTETRCVDGRDITLWDICASNGITHVI 300
DB 1248 VSMFFYNDLVNGTTLQTRLSKLLITASQDPLQPTETRFVDRGAILQWIDIFASNGIITHVI 1307
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIILVTGAVALAAYSYFRINRKTIGFXHF 353
DB 1308 SRPLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSYFRINRRTTIGFQHF 1360

RESULT 3
Q8TES1 PRELIMINARY; PRT; 1736 AA.
AC Q8TES1;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE FLJ00122 protein (Fragment).
GN FLJ00122.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Splice;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074051; BAB84877.1; -.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.

```

```

DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000538; Link.
DR Pfam: PF02469; Fasciclin; 4.
DR Pfam: PF02469; Fasciclin; 4.
DR PRINTS: PRO00193; Xlink; 1.
DR PRINTS: PRO0011; EGFLAMININ.
DR ProDom: PD000918; Link; 1.
DR SMART: SM00181; EGF; 15.
DR SMART: SM00554; FAS1; 5.
DR SMART: SM00445; LINK; 1.
DR PROSITE: PS00022; EGF 1; 4.
DR PROSITE: PS01186; EGF 2; 12.
DR PROSITE: PS0213; FAS1; 5.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 2.
KW EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1736 AA; 187887 MW; 50982047E43925F2 CRC64;

Query Match
Best Local Similarity 95.0%; Score 1772; DB 4; Length 1736;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60
DB 1328 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 1387
QY 61 RSLPGQYKLTFDKAREACANEAAATWATYNQLSYOKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 1388 RSLPGQYKLTFDKAREACANEAAATWATYNQLSYOKAKYHLCSAGWLETGRVAYPTAFAS 1447
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKGVYVGDGFSYGNLQVLMSPF 180
DB 1448 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKGVYVGDGFSYGNLQVLMSPF 1507
QY 181 SLTNFLTEVLAYSNSASARGRAFLHLTDLSTIRGLTFVPSQSLGENETLSGRDIEHHLAN 240
DB 1508 SLTNFLTEVLAYSNSASARGRAFLHLTDLSTIRGLTFVPSQSLGENETLSGRDIEHHLAN 1567
QY 241 VSMFYNDLVNGTTLQTRLSKLLITDQDPLHPTETRCVDRDITLWDICASNGITHVI 300
DB 1568 VSMFYNDLVNGTTLQTRLSKLLITDQDPLHPTETRCVDRDITLWDIFASNGIHI 1627
QY 301 SRXLKAPPAPVTLTHTGLGXGIFXIIILVTGAVALAAYSFRINRKTIGFXHF 353
DB 1628 SRPLKAPPAPVTLTHTGLGAGIFFAILVTGAVALAAYSFRINRKTIGFQHF 1680

RESULT 4
Q8WQ8 PRELIMINARY; PRT; 2551 AA.
AC Q8WQ8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Stablin-2.
GN STAB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Politz O., Gratchev A., McCourt P.A.G., Schledzowski K., Guillot P.,
RA Johansson S., Birk R., Hakiy N., Franke P., Kodelja V., Kannicht C.,
RA "Stablin-1 and Stablin-2 constitute a novel family of fasciclin
RT domain-containing adhesion molecules associated with endothelial-
RT macrophage differentiation and angiogenic processes."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ295695; CAC82105.1; -.
DR GO; GO:0005540; F:hyalurononic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.

```

```

DR InterPro: IPR000782; BIGH3_FAS1.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000538; Link.
DR Pfam: PF02469; Fasciclin; 5.
DR Pfam: PF02469; Fasciclin; 5.
DR PRINTS: PRO00193; Xlink; 1.
DR PRINTS: PRO0011; EGFLAMININ.
DR ProDom: PD000918; Link; 1.
DR SMART: SM00181; EGF; 20.
DR SMART: SM00554; FAS1; 7.
DR SMART: SM00445; LINK; 1.
DR PROSITE: PS00022; EGF 1; 7.
DR PROSITE: PS01186; EGF 2; 16.
DR PROSITE: PS0213; FAS1; 7.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 2.
KW EGF-like domain; Laminin EGF-like domain.
SQ SEQUENCE 2551 AA; 276992 MW; 60A4651CCC2BB69 CRC64;

Query Match
Best Local Similarity 95.0%; Score 1772; DB 4; Length 2551;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60
DB 2143 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 2202
QY 61 RSLPGQYKLTFDKAREACANEAAATWATYNQLSYOKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 2203 RSLPGQYKLTFDKAREACANEAAATWATYNQLSYOKAKYHLCSAGWLETGRVAYPTAFAS 2262
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKGVYVGDGFSYGNLQVLMSPF 180
DB 2263 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKGVYVGDGFSYGNLQVLMSPF 2322
QY 181 SLTNFLTEVLAYSNSASARGRAFLHLTDLSTIRGLTFVPSQSLGENETLSGRDIEHHLAN 240
DB 2323 SLTNFLTEVLAYSNSASARGRAFLHLTDLSTIRGLTFVPSQSLGENETLSGRDIEHHLAN 2382
QY 241 VSMFYNDLVNGTTLQTRLSKLLITDQDPLHPTETRCVDRDITLWDICASNGITHVI 300
DB 2383 VSMFYNDLVNGTTLQTRLSKLLITDQDPLHPTETRCVDRDITLWDIFASNGIHI 2442
QY 301 SRXLKAPPAPVTLTHTGLGXGIFXIIILVTGAVALAAYSFRINRKTIGFXHF 353
DB 2443 SRPLKAPPAPVTLTHTGLGAGIFFAILVTGAVALAAYSFRINRKTIGFQHF 2495

RESULT 5
Q8IUG9 PRELIMINARY; PRT; 2551 AA.
AC Q8IUG9;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE FELE-2.
GN FELE-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Adachi H., Tsujimoto M.;
RA MEDLINE=22206516; PubMed=12077138;
RT "FELE-1, a novel scavenger receptor with in vitro bacteria-binding and
RT angiogenesis-modulating activities."
RL J. Biol. Chem. 277:34264-34270(2002).
DR EMBL; AB052958; BAC15608.1; -.
DR GenBank; HGNC:18629; STAB2.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyalurononic acid binding; IEA.

```

DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR000782; BIGH3_FAS1.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR000538; Link.
 DR InterPro; IPR000737; Squash.
 DR Pfam; PF00008; EGF; 16.
 DR Pfam; PF02469; Fasciclin; 6.
 DR Pfam; PF00193; Xlink; 1.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR ProDom; PD000318; Link; 1.
 DR SMART; SM00181; EGF; 22.
 DR SMART; SM00179; EGF_Ca; 8.
 DR SMART; SM00180; EGF_Lam; 5.
 DR SMART; SM00554; FAS1; 7.
 DR SMART; SM00445; LINK; 1.
 DR SMART; SM00286; PTI; 8.
 DR PROSITE; PS00022; EGF 1; 7.
 DR PROSITE; PS01186; EGF 2; 16.
 DR PROSITE; PS0213; FAS1; 7.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
 DR SEQUENCE 2551 AA; 277026 MW; 3ADDF3F5BD2019C4 CRC64;
 SQ

Query Match 95.0%; Score 1772; DB 4; Length 2551;
 Best Local Similarity 94.6%; Pred. No. 1.2e-157;
 Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
 DB 2143 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 2202

QY 61 RSLPGQYKLTDFKAREACANEATWATYNQLSYQKAKYHLCGAGWLTGRTVAYPTAFAS 120
 DB 2203 RSLPGQYKLTDFKAREACANEATWATYNQLSYQKAKYHLCGAGWLTGRTVAYPTAFAS 2262

QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDFSYSGNLLQVLMSPF 180
 DB 2263 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDFSYSGNLLQVLMSPF 2322

QY 181 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 DB 2323 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 2382

QY 241 VSMFFYNDLVNGTTLTQRLGSKLLITDRODPLHPTETRCVDRDLEWDICASNGITHVI 300
 DB 2383 VSMFFYNDLVNGTTLTQRLGSKLLITDRODPLHPTETRCVDRDLEWDICASNGITHVI 2442

QY 301 SRXLKAPPAPVTLTHTGLGXGIFXXIILVTGAVAAAYSYFRINRKTIGFXHF 353
 DB 2443 SRPLKAPPAPVTLTHTGLGAGIFFAILVTGAVAAAYSYFRINRRTIGFQHF 2495

RESULT 6
 Q7Z5N9 PRELIMINARY; PRT; 2551 AA.
 AC Q7Z5N9;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE FE2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Park S.-Y., Kim I.-S.;
 RT "FE2, a novel cell adhesion molecule of Fas-1 superfamily mediates cell-cell interaction.";

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY311388; AAP74958.1; -
 SQ SEQUENCE 2551 AA; 276986 MW; 3ACB6AC63CB80044 CRC64;

Query Match 95.0%; Score 1772; DB 4; Length 2551;
 Best Local Similarity 94.6%; Pred. No. 1.2e-157;
 Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
 DB 2143 MTGPKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 2202

QY 61 RSLPGQYKLTDFKAREACANEATWATYNQLSYQKAKYHLCGAGWLTGRTVAYPTAFAS 120
 DB 2203 RSLPGQYKLTDFKAREACANEATWATYNQLSYQKAKYHLCGAGWLTGRTVAYPTAFAS 2262

QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDFSYSGNLLQVLMSPF 180
 DB 2263 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDFSYSGNLLQVLMSPF 2322

QY 181 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 DB 2323 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 2382

QY 241 VSMFFYNDLVNGTTLTQRLGSKLLITDRODPLHPTETRCVDRDLEWDICASNGITHVI 300
 DB 2383 VSMFFYNDLVNGTTLTQRLGSKLLITDRODPLHPTETRCVDRDLEWDICASNGITHVI 2442

QY 301 SRXLKAPPAPVTLTHTGLGXGIFXXIILVTGAVAAAYSYFRINRKTIGFXHF 353
 DB 2443 SRPLKAPPAPVTLTHTGLGAGIFFAILVTGAVAAAYSYFRINRRTIGFQHF 2495

RESULT 7
 Q9NRY3 PRELIMINARY; PRT; 897 AA.
 ID Q9NRY3;
 AC Q9NRY3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CD44-like precursor FELL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tao Q., Zhang W., Cao X.;
 RT "Molecular cloning and characterization of human FELL sharing homology with CD44.";
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF160476; AAF82398.1; -
 DR HSSP; P98066; 1TSG.
 DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR000782; BIGH3_FAS1.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR000538; Link.
 DR Pfam; PF00008; EGF; 4.
 DR Pfam; PF02469; Fasciclin; 2.
 DR Pfam; PF00193; Xlink; 1.
 DR PRINTS; PR01265; LINKMODULE.
 DR ProDom; PD000918; Link; 1.
 DR SMART; SM00181; EGF; 5.
 DR SMART; SM00554; FAS1; 3.
 DR SMART; SM00445; LINK; 1.
 DR PROSITE; PS00022; EGF 1; 2.
 DR PROSITE; PS01186; EGF 2; 4.
 DR PROSITE; PS0213; FAS1; 3.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.

```
KW EGF-like domain.
SQ SEQUENCE 897 AA; 97585 MW; EB920AF36101E388 CRC64;

Query Match
Best Local Similarity 94.3%; Score 1769; DB 4; Length 897;
Matches 333; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPKHKCECKSHVYVGDGLNCEPEPOLPIDRCLQNGQCHADAKCVDLHFQDTTGVVPHL 60
DB 489 MTGPKHKCECKSHVYVGDGLNCEPEPOLPIDRCLQNGQCHADAKCVDLHFQDTTGVVPHL 548
QY 61 RSLPGQYKLTFTDKAREACANEAAATWATYNQLSYXKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 549 RSLPGQYKLTFTDKAREACANEAAATWATYNQLSYXKAKYHLCSAGWLETGRVAYPTAFAS 608
QY 121 QNCGSGVWGIYDYGPRPNKSEMDVFCYRMKDVNCTXKGYVGDGFSYSGNLLQVLMSPF 180
DB 609 QNCGSGVWGIYDYGPRPNKSEMDVFCYRMKDVNCTXKGYVGDGFSYSGNLLQVLMSPF 668
QY 181 SUTNLTFLTEVLAYSNSARGAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
DB 669 SUTNLTFLTEVLAYSNSARGAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 728
QY 241 VSMFFYNDLVNGTTLQTRLSKLLITDRODPLHPTETRCVDGRDLEWDICASNGITHVI 300
DB 729 VSMFFYNDLVNGTTLQTRLSKLLITDRODPLHPTETRCVDGRDLEWDICASNGITHVI 788
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGVALAAYSYFRINRKTIGFXHF 353
DB 789 SRPLKAPPAPVTLXHTGLGXGIFXXIILVTGVALAAYSYFRINRKTIGFXHF 841

RESULT 8
Q8CFM6 PRELIMINARY; PRT; 1431 AA.
AC Q8CFM6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hyaluronan receptor for endocytosis HARE precursor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22169209; PubMed=12181351;
RA Zhou B., Weigel J.A., Saxena A., Weigel P.H.;
RT "Molecular Cloning and Functional Expression of the Rat 175-kDa
RL Mol. Biol. Cell 13:2853-2868 (2002).
DR EMBL; AY007370; AAG13634.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF02469; Fasciclin; 3.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00179; EGF CA; 2.
DR SMART; SM00554; FAS1; 4.
DR SMART; SM00445; LINK; 1.
```

```
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS0213; FAS1; 4.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
FT NON_TER.
SQ SEQUENCE 1431 AA; 155914 MW; 3340A8F7AF366722 CRC64;

Query Match
Best Local Similarity 79.3%; Score 1514.5; DB 11; Length 1431;
Matches 280; Conservative 29; Mismatches 43; Indels 1; Gaps 1;

QY 1 MTGPKHKCECKSHVYVGDGLNCEPEPOLPIDRCLQNGQCHADAKCVDLHFQDTTGVVPHL 60
DB 1008 MTGPKHKCECKSHVYVGDVDEPEQLFLDRCLQNGQCHPDASCADLYFQDTTGVVPHL 1067
QY 61 RSLPGQYKLTFTDKAREACANEAAATWATYNQLSYXKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 1068 RSLPGQYKLTFTDKAREACANEAAATWATYNQLSYXKAKYHLCSAGWLETGRVAYPTAFAS 1127
QY 121 QNCGSGVWGIYDYGPRPNKSEMDVFCYRMKDVNCTXKGYVGDGFSYSGNLLQVLMSPF 180
DB 1128 QNCGANVWGIYDYGSRANKSEMDVFCYRMKDVNCTCKAGYVGDGFCSCGNLLQVLMSPF 1187
QY 181 SUTNLTFLTEVLAYSNSARGAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
DB 1188 SUTNLTFLTEVLAYSNSARGAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLTN 1247
QY 241 VSMFFYNDLVNGTTLQTRLSKLLITDRODPLHPTETRCVDGRDLEWDICASNGITHVI 300
DB 1248 VNVSYNDLVNGTFLRTMLGSQLITPQDQLH-QETRFVDSRILQMDIIRANGILHII 1306
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGVALAAYSYFRINRKTIGFXHF 353
DB 1307 SEPLRAPPTAATAAHSGLGTGIFCAVLVLTGAIALAAYSYFRILKQRTTGFQRF 1359

RESULT 9
Q8BM87 PRELIMINARY; PRT; 894 AA.
AC Q8BM87;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Similar to CD44-like precursor FELL (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK034522; BAC28741.1; -.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF02469; Fasciclin; 2.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 5.
```



```

DR SMART; SM00180; EGF Lam; 2.
DR SMART; SM00554; FAS1; 3.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS02013; FAS1; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
FT NON TER 1
SQ SEQUENCE 894 AA; 97059 MW; 3286E2A223CFB96 CRC64;

Query Match 79.7%; Score 1487; DB 11; Length 894;
Best Local Similarity 77.9%; Pred. No. 2.3e-131;
Matches 275; Conservative 32; Mismatches 44; Indels 2; Gaps 2;

QY 1 MTGPKKCKCKSHYVGDGLNCEPEQLPIDRCLDNGQCHADAKVDLHFQDTTVGVFHL 60
Db 486 MTGPKKCKCKSHYVGDGRDCEPEQLPIDRCLDNGQCHPDANCVDLHFQDTTVGVFHL 545

QY 61 RSPGLQYKLTDFDKAREACANEATMATYNQLSYXOKAKYHLCAGWLETGRVAYPTAFAS 120
Db 546 RSPGLQYKLTDFDKAREACAEAASTATYNQLSYAQAKYHLCAGWLESGRVAYPTIYAS 605

QY 121 QNCSGGVVGIYDYGPRPNKSEMDVFCYRMKDVNCTXXKVGVDGFSYGNLLQVLMSEFP 180
Db 606 KKC-ANIYVIGIYDYGTRTNKSEMDVFCYRMKDVNCTCKAGYVGDGFCNGNLLQVLMSEFP 664

QY 181 SLTNFLTEVLAYSNSARGRAFLKHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLN 240
Db 665 SLTNFLTEVLVFSRSSAQGRAFLKHLTDLISIRGTLFVPQNSGLPKNKSLSGRDIEHHLN 724

QY 241 VSMFYNDLVNGTTLQTRLGSKLLITDRQDPLHPHTETRCVDGRDTELDWDCASNGITHI 300
Db 725 VNVSYFDLVNGTTLQTRLGSKLLITSSQDLH-QEARFVDRGAILQWDIIASNGVLHII 783

QY 301 SRXLKAPAPVTLXHTGLGXGIFXXIIIVTCAVALAAYSYERINRKTIGEXHF 353
Db 784 SEPLKAPPTAATAHSGLTGIFCAVLVTGAILAALAYSYFLNRQTTGFRFR 836

RESULT 10
Q8R4U0
ID Q8R4U0 PRELIMINARY; PRT; 2559 AA.
AC Q8R4U0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Stabilin-2.
GN STAB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=1818465; PubMed=11829752;
RX Politz O., Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,
RA Johansson S., Svineng G., Franke P., Kannicht C., Kzyshkowska J.,
RA Longati P., Velten F.W., Johansson S., Goerd S.;
RT "Stabilin-1 and -2 constitute a novel family of fasciclin-like
hyaluronan receptor homologues.";
RL Biochem J 362:155-164(2002).
DR EMBL; AF364951; AAL9184.2; -.
DR MGD; MGI:2178743; Stab2.
DR GO; GO:0003779; F.actin binding; IEA.
DR GO; GO:0005540; F.hyaluronic acid binding; IEA.
DR GO; GO:0005198; P.structural molecule activity; IEA.
DR GO; GO:0007155; P.cell adhesion; IEA.
DR InterPro; IPR001589; Actbind actnin.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.

```

```

DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 13.
DR Pfam; PF02469; Fasciclin; 6.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRODOM; PD000918; Link; 1.
DR SMART; SM00181; EGF; 21.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS02013; FAS1; 7.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW EGF-like domain.
SQ SEQUENCE 2559 AA; 277530 MW; 1C9855AD61EFF015 CRC64;

Query Match 79.7%; Score 1487; DB 11; Length 2559;
Best Local Similarity 77.9%; Pred. No. 1e-130;
Matches 275; Conservative 32; Mismatches 44; Indels 2; Gaps 2;

QY 1 MTGPKKCKCKSHYVGDGLNCEPEQLPIDRCLDNGQCHADAKVDLHFQDTTVGVFHL 60
Db 2151 MTGPKKCKCKSHYVGDGRDCEPEQLPIDRCLDNGQCHPDANCVDLHFQDTTVGVFHL 2210

QY 61 RSPGLQYKLTDFDKAREACANEATMATYNQLSYXOKAKYHLCAGWLETGRVAYPTAFAS 120
Db 2211 RSPGLQYKLTDFDKAREACAEAASTATYNQLSYAQAKYHLCAGWLESGRVAYPTIYAS 2270

QY 121 QNCSGGVVGIYDYGPRPNKSEMDVFCYRMKDVNCTXXKVGVDGFSYGNLLQVLMSEFP 180
Db 2271 KKC-ANIYVIGIYDYGTRTNKSEMDVFCYRMKDVNCTCKAGYVGDGFCNGNLLQVLMSEFP 2329

QY 181 SLTNFLTEVLAYSNSARGRAFLKHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLN 240
Db 2330 SLTNFLTEVLVFSRSSAQGRAFLKHLTDLISIRGTLFVPQNSGLPKNKSLSGRDIEHHLN 2389

QY 241 VSMFYNDLVNGTTLQTRLGSKLLITDRQDPLHPHTETRCVDGRDTELDWDCASNGITHI 300
Db 2390 VNVSYFDLVNGTTLQTRLGSKLLITSSQDLH-QEARFVDRGAILQWDIIASNGVLHII 2448

QY 301 SRXLKAPAPVTLXHTGLGXGIFXXIIIVTCAVALAAYSYERINRKTIGEXHF 353
Db 2449 SEPLKAPPTAATAHSGLTGIFCAVLVTGAILAALAYSYFLNRQTTGFRFR 2501

RESULT 11
Q8UF98
ID Q8UF98 PRELIMINARY; PRT; 1069 AA.
AC Q8UF98;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKZP2434E0321.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133021; CAB61358.1; -.
DR PIR; T42681; T42681.
DR HSP; P98066; TRSG.
DR GO; GO:0005540; F.hyaluronic acid binding; IEA.
DR GO; GO:0005198; P.structural molecule activity; IEA.
DR GO; GO:0007155; P.cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.

```

```

DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 7.
DR Pfam; PF02469; Fasciclin; 2.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 8.
DR SMART; SM00554; FAS1; 3.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS0213; FAS1; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
DR Hypothetical protein; EGF-like domain; Laminin EGF-like domain.
KW NON_TER
SQ SEQUENCE 1069 AA; 115666 MW; 04B0960932164D63 CRC64;

Query Match 79.4%; Score 1480; DB 4; Length 1069;
Best Local Similarity 81.0%; Pred. No. 1.4e-130;
Matches 286; Conservative 4; Mismatches 27; Indels 36; Gaps 2;

QY 1 MTGPKHKCECKSHVYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDTTVGVFHL 60
Db 697 MTGPKHKCECKSHVYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDTTVGVFHL 756

QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSYQKAKYHLCSAGWLETGRVAYPTAFAS 120
Db 757 RSPGQYKLTDFDKAREACANEATMATYNQLSYQKAKYHLCSAGWLETGRVAYPTAFAS 816

QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVCNTXXKVYVGDGFSYSGNLLQVLMSPF 180
Db 817 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKG-----SAGLFQQLSRP 860

QY 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENITSLGRDIEHILAN 240
Db 861 CLS-----RFPDDLISIRGTLFVPQNSGLGENITSLGRDIEHILAN 900

QY 241 VSMFYNDLVNGTTIQLRGLSKLLITDRQDPLHPTETRCVDCGRDLEWDICASNGITHVI 300
Db 901 VSMFYNDLVNGTTIQLRGLSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIHHVI 960

QY 301 SRXLKAPAPVTLXTGLGXGIFXXIILVTGAVALAAASYFRINRKTIGFQHF 353
Db 961 SRPLKAPAPVTLTGLGAGIFFAILIIVTGAVALAAASYFRINRRTIGFQHF 1013

RESULT 12
Q8K0K6 PRELIMINARY; PRT; 1238 AA.
AC Q8K0K6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to stabilin 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031166; AAH31166.1;
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BigH3_FAS1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.

```

```

DR InterPro; IPR000538; Link.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00008; EGF; 8.
DR Pfam; PF02469; Fasciclin; 2.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00554; FAS1; 3.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 4.
DR PROSITE; PS01186; EGF_2; 10.
DR PROSITE; PS0213; FAS1; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
DR PROSITE; PS01241; LINK; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
KW EGF-like domain.
KW NON_TER
SQ SEQUENCE 1238 AA; 132618 MW; C101FB9F26470FCD CRC64;

Query Match 35.6%; Score 664.5; DB 11; Length 1238;
Best Local Similarity 39.9%; Pred. No. 1.e-53;
Matches 146; Conservative 48; Mismatches 145; Indels 27; Gaps 8;

QY 2 TGPQKHKECKSHVYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDTTVGVFHL 60
Db 820 TGPNTTRCECHVYVGDGLQCLELEPPVDRCLGSSPCHTDALCTDLHFQEKQAGVEHI 879

QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSYQKAKYHLCSAGWLETGRVAYPTAFAS 120
Db 880 QATGPGYGLTSEAKACEGQGVASLPQSAQQGLGHVCFVGNLANGSAAHVPVTPA 939

QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVCNTXXKVYVGDGFSYSGNLLQVLMSPF 179
Db 940 ADCGNRVGVVSLGVRKNLSLWDAVCYRVQDVACQCRAGFVGDGISTCNGLLDVLAAT 999

QY 180 PSNLNFLTVELAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENITSLGRDIEHILA 239
Db 1000 ANFTFYGMLLGVANATQGLFEMDFLEDELTYSKTLFVFNKGFVNDNLTSGPDLEHAS 1059

QY 240 NVSMFFYNDLVNGTTIQLRGLSKLLITDR---QDPLHPTETRCVDCGRDLEWDICASNGI 296
Db 1060 NATFELSIN-ASRGTLTPAHSGLSLFSIDTGPNTSLVLPAGVAVVSHVIVMDIMAFNGI 1118

QY 297 THVISRXLKAP-----PAPVTLXHTGLGXGIFXXIILVTGAVALAAASYFRIN 344
Db 1119 IHALASPLLMPPQTRAVLGSPPPPVAL--SLGVVVTSGTLGLVAGAL-----YLGR 1169

QY 345 RKTIGF 350
Db 1170 GKPPGF 1175

RESULT 13
Q8R4Y4 PRELIMINARY; PRT; 2571 AA.
AC Q8R4Y4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Stabilin-1.
GN STAB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=21818465; PubMed=11829752;
RA Politz O., Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,
RA Johansson S., Svineng G., Franke P., Kannicht C., Kzyshkowska J.,

```

RESULT 14	
Q8IUH1	
ID	Q8IUH1
AC	Q8IUH1
DT	01-MJ

Db 2352 ANFTFYGMILLGYANATQGLDFLDDELTUKLFPVNEGFVDMNTLSGPDLELHAS 2411
QY 240 NVSMFFYNLDLVNGTTLQTRGLSKILLITDR---QDPLHPTETRCVDGRDITLEWDICASNGI 296
Db 2412 NATILSAN-ASQKLLPAHSGLSLIIISDAGPDNSSWAPVPGTVVVSRIIVWDIMAFNGI 2470
QY 297 THVISRXLKAPPAPVTLL-----XHTGLGXGIFXXIILVTGAV--ALAAYSYPFRINRKT 347
Db 2471 IHALASPLAPPQPOQAVLAPEAPPAAGVGA-----VLAAGALLGLVAGALYLRARGKP 2524
QY 348 IGF 350
Db 2525 MGF 2527

Search completed: March 30, 2004, 15:42:31
Job time : 48 secs

Db 2506 MGF 2508
RESULT 15
Q93072 PRELIMINARY; PRT; 2589 AA.
ID Q93072
AC Q93072
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein KIAA0246 (Fragment).
GN KIAA0246.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayasi Y., Ohara O.,
RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT the coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).
DR EMBL; D87433; BAA13377.2; -.
DR HSSP; P98066; 1TSG.
DR Genew; HGNC:18628; STAB1.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 12.
DR Pfam; PF02469; Fasciclin; 4.
DR Pfam; PF00193; Xlink; 1.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 25.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 2.
DR PROSITE; PS00022; EGF 1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS0213; FAS1; 7.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
DR PROSITE; PS01241; LINK; 1.
KW Hypothetical protein; EGF-like domain.
FT NON TER 1
SQ SEQUENCE 2589 AA; 277512 MW; 86F996423001C756 CRC64;
Query Match 35.5%; Score 661.5; DB 4; Length 2589;
Best Local Similarity 39.9%; Pred. No. 1.1e-52;
Matches 145; Conservative 49; Mismatches 148; Indels 21; Gaps 7;
QY 2 TGPGKHCKECKSHYVGDGLNC-EPQLPIDRCLQDNGQCHADAKVDLHFQDITTVGVFHL 60
Db 2172 TGLNTRCECHAGYVGDGLQCLESEPPVDRCLGPPCHSDAMCTDLHFQEKRAGVFHL 2231
QY 61 RSPFGQYKLTDFKAREACANEAAATWATYNQLSYQKAKYHLCAGMLETGRVAYPTAFAS 120
Db 2232 QATSGPYGLNFSEAEACEAQAQAVLASFPQLSAAQQLGHLCLMGWLANGSTAHFVWFPV 2291
QY 121 QNCGSGVGVIVDYGPRPNKSEWDFCYRMKDVNCTYKVGVDGFS-YSGNLLQVLMNF 179
Db 2292 ADCNGRVRGIVSLGARKNLSEWDAYCFRQDVACRCRNGFVGDTSTCNGKLLDVLAAT 2351
QY 180 PSLTFLTEVLAYNSSARGRAFLHLDLSIRGTILFVPQNSGLGENETLSGRDIEHHLA 239

No ant

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 30, 2004, 15:34:05 ; Search time 17 Seconds
(without alignments)
1081.221 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPKHKCKESHVVGDL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	264	14.2	275	1	TSG6_MOUSE
2	264	14.2	276	1	TSG6_MOUSE
3	262	14.0	277	1	TSG6_MOUSE
4	175	9.4	2109	1	PGCA_CHICK
5	174	9.3	2333	1	PGCA_MOUSE
6	171	9.2	2124	1	PGCA_MOUSE
7	170	9.1	2364	1	PGCA_MOUSE
8	169.5	9.1	354	1	PLK_MOUSE
9	169.5	9.1	356	1	PLK_MOUSE
10	166.5	8.9	354	1	PLK_MOUSE
11	166	8.9	340	1	PLK_MOUSE
12	165.5	8.9	354	1	PLK_MOUSE
13	165.5	8.9	355	1	PLK_MOUSE
14	165	8.8	2132	1	PGCA_MOUSE
15	164	8.8	341	1	BRAI_MOUSE
16	164	8.8	341	1	BRAI_MOUSE
17	163.5	8.8	354	1	PLK_MOUSE
18	162	8.7	1321	1	PGCN_MOUSE
19	162	8.7	2415	1	PGCA_MOUSE
20	161.5	8.7	354	1	PLK_MOUSE
21	158.5	8.5	1268	1	PGCN_MOUSE
22	158	8.5	537	1	PGCA_MOUSE
23	157.5	8.4	883	1	PGCB_MOUSE
24	157.5	8.4	1257	1	PGCN_MOUSE
25	155.5	8.3	394	1	PGCA_MOUSE
26	155.5	8.3	912	1	PGCB_MOUSE
27	154.5	8.3	862	1	PGCV_MOUSE
28	154.5	8.3	3562	1	PGCV_MOUSE
29	153	8.2	2738	1	PGCV_MOUSE
30	152.5	8.2	3358	1	PGCV_MOUSE
31	151.5	8.1	3381	1	PGCV_MOUSE
32	150.5	8.1	883	1	PGCB_MOUSE
33	150.5	8.1	3396	1	PGCV_MOUSE

34	147	7.9	417	1	PGCB_MOUSE
35	109.5	5.9	362	1	CD44_MOUSE
36	107.5	5.8	431	1	CD44_MOUSE
37	105	5.6	810	1	NEL1_MOUSE
38	104.5	5.6	359	1	CD44_MOUSE
39	101.5	5.4	1504	1	SLIT_MOUSE
40	100	5.4	2907	1	PN2_MOUSE
41	99	5.3	778	1	CD44_MOUSE
42	98	5.3	742	1	CD44_MOUSE
43	97.5	5.2	503	1	CD44_MOUSE
44	97	5.2	351	1	CD44_MOUSE
45	97	5.2	2911	1	PN2_MOUSE

ALIGNMENTS

RESULT 1
TSG6_MOUSE
ID TSG6_MOUSE STANDARD; PRT; 275 AA.
AC O08859;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor-inducible protein TSG-6 precursor (TNF-
stimulated gene 6 protein).
GN TNFAIP6 OR TNFIP6 OR TSG6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RC STRAIN=CD-1; TISSUE=Cumulus cell, Embryo, and Oocyte;
RX MEDLINE=98087423; PubMed=9427551;
RA Fuelleop C., Kamath R.V., Li Y., Otto J.M., Salustri A., Olsen B.R.,
Glant T.T., Hascall V.C.;
RT "Coding sequence, exon-intron structure and chromosomal localization
of murine TNF-stimulated gene 6 that is specifically expressed by
expanding cumulus cell-oocyte complexes.";
RL Gene 202:95-102(1997).
CC -!- FUNCTION: Possibly involved in cell-cell and cell-matrix
interactions during inflammation and tumorigenesis (By
similarity).
CC -!- DEVELOPMENTAL STAGE: Expressed in cumulus cell-oocyte complexes
during expansion in vivo.
CC -!- SIMILARITY: Contains 1 link domain.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; U83903; AAC53527.1; -.
CC PIR; JC6506; JC6506.
CC HSP; P98066; 1TSG.
CC MGD; MGI:1195266; Tnfalp6.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000538; Link.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00193; Xlink; 1.
CC PRINTS; PR01265; LINKMODULE.
CC ProDom; PD000918; Link; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00445; LINK; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS01241; LINK; 1.
CC Cell adhesion; Signal; Glycoprotein.
CC SIGNAL 1 17
CC POTENTIAL.

```

RP ERRATUM.
RX MEDLINE=94012707; PubMed=8407990;
RA Feng P., Liaw G.;
RL J. Biol. Chem. 268:21453-21453(1993).
CC -I- FUNCTION: Possibly involved in cell-cell and cell-matrix
CC interactions during inflammation and tumorigenesis.
CC -I- TISSUE SPECIFICITY: Vascular smooth muscle cells.
CC -I- DEVELOPMENTAL STAGE: Fetal.
CC -I- INDUCTION: By serum and growth factor.
CC -I- SIMILARITY: Contains 1 link domain.
CC -I- SIMILARITY: Contains 1 CUB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M86381; AAA03342.1; -.
CC PIR; A48055; A47290.
CC HSP; P98066; LTSG.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000538; Link.

```

DR	Pfam; PF00431; CUB; 1.
DR	Pfam; PF00193; Xlink; 1.
DR	PRINTS; PR01265; LINKMODULE.
DR	ProDom; PD000918; Link; 1.
DR	SMART; SM00042; CUB; 1.
DR	SMART; SM00445; LINK; 1.
DR	PROSITE; PS01180; CUB; 1.
DR	PROSITE; PS01241; LINK; 1.
KW	Cell adhesion; Signal; Glycoprotein.
FT	SIGNAL 1 19 POTENTIAL.
FT	CHAIN 20 276 TUMOR NECROSIS FACTOR-INDUCIBLE PROTEIN
FT	TSG-6. LINK.
FT	DOMAIN 53 128 LINK.
FT	DOMAIN 135 247 CUB.
FT	DISULFID 58 127 BY SIMILARITY.
FT	DISULFID 82 103 BY SIMILARITY.
FT	DISULFID 135 161 BY SIMILARITY.
FT	DISULFID 188 210 BY SIMILARITY.
FT	CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE 276 AA; 31081 MW; 3BDCSD9A24B2F75A CRC64;
Query Match 14.2%; Score 264; DB 1; Length 276;	
Best Local Similarity 43.3%; Pred. NO. 4.7e-16;	
Matches 45; Conservative 16; Mismatches 43; Indels 0; Gaps	
QY	52 DTTVGVFHLRSLPGVQLNFDKAREACANEATMATYNOLSYXQAKYHLCGAGWLETGR
Db	32 EQAGVYHRASRGKYLYIAKAACEPEGGRLATYKQLEARKTGIFHVCAAGWMAKGR
QY	112 VAYTAFASQNCGSGVGVGDVGPENPKSEMDVFCYRMKDVCNC 155
Db	92 VGYPVKPSNGCFGKTGIIDYGIRLNRSERWDACYNPFAKEC 135
RESULT 3	
TSG6 HUMAN	STANDARD; PRT; 277 AA.
ID	TSG6_HUMAN
AC	P80066; Q8WWI9;
DT	01-FEB-1996 (Rel. 33, Created)
DC	01-FEB-1996 (Rel. 33, Last sequence update)
DD	15-MAR-2004 (Rel. 43, Last annotation update)
DE	tumor necrosis factor-inducible protein TSG-6 precursor (TNF-
DN	stimulated gene 6 protein) (Hyaluronate-binding protein).
GN	TNFAIP6 OR TSG6.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RX	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Fibroblast;
RA	MEDLINE=11854277;
RX	Lee T.H., Wisniewski H.-G., Vilcek J.;
RT	"A novel secretory tumor necrosis factor-inducible protein (TSG-6) is
RT	a member of the family of hyaluronate binding proteins, closely
RT	related to the adhesion receptor CD44.";
RL	J. Cell Biol. 116:545-557(1992).
RN	[2]
RP	SEQUENCE FROM N.A., AND VARIANT GLN-144.
RX	MEDLINE=21975206; PubMed=11854277;
RA	Nenwich H.A., Mustafa Z., Rugg M.S., Marsden B.D., Cordell M.R.,
RA	Mahtoney D.J., Jenkins S.C., Dowling B., Fries E., Milner C.M.,
RA	Loughlin J., Day A.J.;
RT	"A novel allelic variant of the human TSG-6 gene encoding an amino
RT	acid difference in the CUB module. Chromosomal localization,
RL	frequency analysis, modeling, and expression.";
RL	J. Biol. Chem. 277:15354-15362(2002).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung, and Spleen;
RA	MEDLINE=22388257; PubMed=12477932;
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.;


```

RX MEDLINE=94107258; PubMed=8280087;
RA Chandrasekaran L., Tanzer M.L.;
RL Biochem. J. 296:885-887(1993).
RN [5]
RP SEQUENCE OF 1492-1610 FROM N.A.
RC STRAIN=White leghorn; TISSUE=Chondrocytes;
RX MEDLINE=95128519; PubMed=7827752;
RA Primorac D., Stover M.L., Clark S.H., Rowe D.W.;
RT "Molecular basis of nanomelia, a heritable chondrodystrophy of
RT chicken.";
RL Matrix Biol. 14:297-305(1994).
RN [6]
RP SEQUENCE OF 1894-2109 FROM N.A.
RX MEDLINE=89008500; PubMed=3170613;
RA Tanaka T., Har-El R., Tanzer M.L.;
RT "Partial structure of the gene for chicken cartilage proteoglycan
RT core protein.";
RL J. Biol. Chem. 263:15831-15835(1988).
RN [7]
RP SEQUENCE OF 1693-1855 AND 1893-2109 FROM N.A.
RX MEDLINE=86259736; PubMed=3460082;
RA Sai S., Tanaka T., Koshier R.A., Tanzer M.L.;
RT "Cloning and sequence analysis of a partial cDNA for chicken
RT cartilage proteoglycan core protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5081-5085(1986).
CC -!- FUNCTION: This proteoglycan is a major component of extracellular
CC matrix of cartilaginous tissues. A major function of this protein
CC is to resist compression in cartilage. It binds avidly to
CC hyaluronic acid via an amino-terminal globular region. May play a
CC regulatory role in the matrix assembly of the cartilage.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1;
CC IsoId=P07898-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P07898-2; Sequence=VSP_003073;
CC -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC terminus of the proteoglycan, while another globular region, G3,
CC makes up the COOH terminus. G1 contains link domains and thus
CC consists of three disulfide-bonded loop structures designated as
CC the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC and the chondroitin sulfate (CS) attachment domains lie between G2
CC and G3.
CC -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC chains, N-linked and O-linked oligosaccharides.
CC -!- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF NANOMELIA, A LETHAL
CC CONNECTIVE TISSUE DISORDER AFFECTING CARTILAGE DEVELOPMENT
CC (CHONDRODYSPLASIA) CHARACTERIZED BY SHORTENED AND MALFORMED LIMBS.
CC AGGREGAN IS TRUNCATED AT ITS C-TERMINAL IN THE CS-2 BINDING DOMAIN
CC AND IS NOT ANMORE SECRETED FROM THE CHONDROCYTES.
CC -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; I21913; AAB19128.1; -.
DR EMBL; M38187; AAA48731.1; -.
DR EMBL; M88101; -. NOT ANNOTATED_CDS.
DR EMBL; S74657; AAC60751.1; -.
DR EMBL; S74656; AAC60751.1; JOINED.
DR EMBL; J04028; AAA48719.1; -.

```

```

DR EMBL; M13993; AAA48720.1; -.
DR PIR; I50421; I50421.
DR HSP; P08709; IBF9.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000152; Asx_hydroxyl_s.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXSG.
DR InterPro; IPR000436; Sushi_SCR_OCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; Lectin_c; 1.
DR Pfam; PF02339; SGXSG; 56.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLSCT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_Like; 1.
DR PROSITE; PS01241; LINK; 4.
KW Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW Alternative splicing; Repeat; Immunoglobulin domain.
FT CHAIN 1 16
FT DOMAIN 17 2109 AGGREGAN CORE PROTEIN.
FT DOMAIN 34 143 IG-LIKE V-TYPE.
FT DOMAIN 166 243 LINK 1.
FT DOMAIN 264 346 LINK 2.
FT DOMAIN 537 614 LINK 3.
FT DOMAIN 635 716 LINK 4.
FT DOMAIN 1363 1742 19 X 20 AA TANDEM-REPEAT.
FT DOMAIN 1855 1892 EGF-LIKE.
FT DOMAIN 1901 2019 C-TYPE LECTIN.
FT DOMAIN 2023 2081 SUSHI.
FT DOMAIN 48 137 G1-A.
FT DOMAIN 148 243 G1-B.
FT DOMAIN 249 346 G1-B'.
FT DOMAIN 519 613 G2-B.
FT DOMAIN 620 715 G2-B'.
FT DOMAIN 718 803 KS.
FT DOMAIN 805 1264 CS-1.
FT DOMAIN 1265 1742 CS-2.
FT DOMAIN 1893 2109 G3.
FT DISULFID 51 129 BY SIMILARITY.
FT DISULFID 171 242 BY SIMILARITY.
FT DISULFID 195 216 BY SIMILARITY.
FT DISULFID 269 345 BY SIMILARITY.
FT DISULFID 293 314 BY SIMILARITY.
FT DISULFID 542 613 BY SIMILARITY.
FT DISULFID 566 587 BY SIMILARITY.
FT DISULFID 640 715 BY SIMILARITY.
FT DISULFID 664 685 BY SIMILARITY.
FT DISULFID 1859 1870 BY SIMILARITY.
FT DISULFID 1864 1879 BY SIMILARITY.
FT DISULFID 1881 1890 BY SIMILARITY.
FT DISULFID 1897 1908 BY SIMILARITY.
FT DISULFID 1925 2017 BY SIMILARITY.
FT DISULFID 1933 2009 BY SIMILARITY.

```



```

FT DISULFID 2024 2067 BY SIMILARITY.
FT DISULFID 2053 2080 BY SIMILARITY.
FT CARBOHYD 76 76 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 700 700 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 801 801 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1856 1892 Missing (in isoform 2).
FT CONFLICT 362 362 /FTID=VSP_003073.
Query Match 9.4%; Score 175; DB 1; Length 2109;
Best Local Similarity 23.4%; Pred. No. 4.8e-07;
Matches 75; Conservative 31; Mismatches 116; Indels 98; Gaps 12;

QY 10 ECKSHVVDGLNCEPEQIPIDRLQDNGQCHA-----DAKVDLHFQDTTVG 56
DQ 194 QCDAGLADQTVRYIHLPRERCYGDDEFPQVRYTGVRETDTYDVVYCYAEQMOGK--- 250
QY 57 VPHLSPLGQYLTRDKAREACANEAATMATYNQLSYXQAKYHLCSAGWLETGEVAYPT 116
DQ 251 VYATSP---EKFTFOEAFDKCHSGARLATTTGELYLAWKQMDMCSAGWLADRSVRYP 307
QY 117 APASQNGSGVVGVI--VDYGPB-----PNKSMWDFVFCYRMKDVCN-----TKRVG-YV 162
DQ 308 SRAPNCGNLGVRTVYLNPNANQTYGPHPSRYDAICYSGGDFEALVPLGFTDVGTEL 367
QY 163 GDGFSY-----SGNLLQVLSFPPLSLNPL 186
DQ 368 GSAPTQITQTVTEVELPLPRNVTEBEARGSTATLEPMEITATATLYEAFVLPDL--FA 425
QY 187 TEVLAYSNSARGRAFLHLDL-----SIRGTLFVPCNSGLGE----- 225
DQ 426 TSVVETASPRENVTVREITGNAVPEVTVTSVGTAF---TTGMAEVSSVEEAIAVTA 482
QY 226 ---NETLSGRDIEHHLANSV 242
DQ 483 TPLGESASPFTHEDHLVQVT 502

RESULT 5
ID PGCA_CANFA STANDARD; PRT: 2333 AA.
AC Q28343; Q28310;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).
DN AGC1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Glant T.T., Adams M.E., Kwok S.X.F., Huang D., Fuloop C.;
RT "Complete coding sequence and deduced amino acid sequence of aggrecan of canine cartilage.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 774-833 FROM N.A.
RC TISSUE=Cartilage;
EX MEDLINE=95128522; PubMed=7827755;
RA Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT "Length variation in the keratan sulfate domain of mammalian aggrecan.";
RL Matrix Biol. 14:323-328(1994).
RN [3]

```

```

RP SEQUENCE OF 1830-2333 FROM N.A.
RA Adams M.E., Kwok S.X.F., Huang D., Glant T.T., Fuloop C.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 2082-2118 FROM N.A.
RC TISSUE=Cartilage; PubMed=8349621;
RX MEDLINE=93352525;
RA Fuloop C., Walcz E., Vallyon M., Glant T.T.;
RT "Expression of alternatively spliced epidermal growth factor-like domains in aggrecans of different species. Evidence for a novel module.";
RT J. Biol. Chem. 268:17377-17383(1993).
CC -!- FUNCTION: This proteoglycan is a major component of extracellular matrix of cartilaginous tissues. A major function of this protein is to resist compression in cartilage. It binds avidly to hyaluronic acid via an amino-terminal globular region. May play a regulatory role in the matrix assembly of the cartilage.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By similarity).
CC -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino terminus of the proteoglycan, while another globular region, G3, makes up the COOH terminus. G1 contains link domains and thus consists of three disulfide-bonded loop structures designated as the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS) and the chondroitin sulfate (CS) attachment domains lie between G2 and G3.
CC -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate chains, N-linked and O-linked oligosaccharides (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
EMBL; U65989; AAB06238.2; -.
DR EMBL; S74662; AAC60527.1; -.
DR EMBL; L07054; -. NOT_ANNOTATED_CDS.
DR PIR; I46998; I46998.
DR HSSP; P08709; 1BF9.
DR InterPro; IPR0002353; AntifreezeII.
DR InterPro; IPR000152; Asx_hydroxyl_s.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_C; 1.
DR Pfam; PF02339; SGXSG; 66.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.
DR PRINTS; PR0356; ANTIFREEZEII.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_Ca; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.

```

4


```

CC Name=1;
CC IsoId=P13608-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P13608-2; Sequence=VSP_003072;
CC -I- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC terminus of the proteoglycan, while another globular region, G3,
CC makes up the C-terminus. G1 contains link domains and thus
CC consists of three disulfide-bonded loop structures designated as
CC the A, B, and C motifs. G2 is similar to G1. The keratan sulfate (KS)
CC and the chondroitin sulfate (CS) attachment domains lie between G2
CC and G3.
CC -I- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-
CC LINKED (ABOUT 40) OLIGOSACCHARIDES.
CC -I- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN
CC ADULT AND FETAL BOVINE PROTEOGLYCAN.
CC -I- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -I- SIMILARITY: Contains 4 link domains.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 C-type lectin family domain.
CC -I- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -I- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; U76615; AAB38524.1; --
CC ENBL; L07053; -; NOT_ANNOTATED_CDS.
CC PIR; A34234; A39808.
CC PIR; T42630; T42630.
CC DR HSP; P08709; IBF9.
CC DR InterPro; IPRO02353; Antifreeze1.
CC DR InterPro; IPRO00152; Asx_hydroxyl_s.
CC DR InterPro; IPRO00742; EGF_2.
CC DR InterPro; IPRO01881; EGF_Ca.
CC DR InterPro; IPRO06209; EGF_Like.
CC DR InterPro; IPRO07110; Ig_Like.
CC DR InterPro; IPRO03006; Ig_MHC.
CC DR InterPro; IPRO01304; Lectin_C.
CC DR InterPro; IPRO00538; Link.
CC DR InterPro; IPRO03324; SGXSG.
CC DR InterPro; IPRO00436; Sushi_SCR_CCP.
CC DR Pfam; PF00008; EGF; 1.
CC DR Pfam; PF00047; ig; 1.
CC DR Pfam; PF00059; lectin_c; 1.
CC DR Pfam; PF02339; SGXSG; 61.
CC DR Pfam; PF00084; sushi; 1.
CC DR Pfam; PF00193; Xlink; 4.
CC DR PRINTS; PR00356; ANTIFREEZE11.
CC DR PRINTS; PR01265; LINKMODULE.
CC DR ProDom; PD000918; Link; 4.
CC DR SMART; SM00032; CCP; 1.
CC DR SMART; SM00034; CLECT; 1.
CC DR SMART; SM00179; EGF_CA; 1.
CC DR SMART; SM00445; LINK; 4.
CC DR PROSITE; PS00010; ASX_HYDROXYL; 1.
CC DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS50026; EGF_3; 1.
CC DR PROSITE; PS01187; EGF_CA; 1.
CC DR PROSITE; PS50835; IG_LIKE; 1.
CC DR PROSITE; PS00290; IG_MHC; 1.
CC DR PROSITE; PS01241; LINK; 4.
CC KW Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
CC Calcium; Alternative splicing; Repeat; Immunoglobulin domain.
CC SIGNAL 1 16 POTENTIAL.
CC FT CHAIN 17 2364 AGGECAN CORE PROTEIN.
CC FT DOMAIN 25 147 IG-LIKE V-TYPE.
CC FT

```

```

RT structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:3761-3765 (1986).
RN [3]
RX MEDLINE=86140139; PubMed=2419334;
RA Neame P.J., Christner J.E., Baker J.R.;
RT "The primary structure of link protein from rat chondrosarcoma
RT proteoglycan aggregate.";
RL J. Biol. Chem. 261:3519-3535 (1986).
CC -!- FUNCTION: STABILIZE THE AGGREGATES OF PROTEOGLYCAN MONOMERS WITH
CC HYALURONIC ACID IN THE EXTRACELLULAR CARTILAGE MATRIX.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P03994-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P03994-2; Sequence=VSP 005301, VSP 005302;
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22340; AAA41535.1; -
DR EMBL; M22336; AAA41535.1; JOINED.
DR EMBL; M22337; AAA41535.1; JOINED.
DR EMBL; M22339; AAA41535.1; JOINED.
DR EMBL; M22340; AAA41536.1; -
DR EMBL; M22336; AAA41536.1; JOINED.
DR EMBL; M22338; AAA41536.1; JOINED.
DR EMBL; M22339; AAA41536.1; JOINED.
DR PIR; A28654; LKRT2.
DR HSP; P98066; LTSG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG.V.
DR InterPro; IPR000538; Link.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000518; Link; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Connective tissue; Extracellular matrix;
KW Hyaluronic acid; Proteoglycan; Immunoglobulin domain; Repeat;
KW Alternative splicing.
FT PROPEP 1 9
FT CHAIN 10 354 BY SIMILARITY.
FT DOMAIN 38 152 PROTEOGLYCAN LINK PROTEIN.
FT DOMAIN 176 253 IG-LIKE V-TYPE.
FT DOMAIN 274 350 LINK 1.
FT DISULFID 61 139 LINK 2.
FT DISULFID 181 252
FT DISULFID 205 226
FT DISULFID 279 349
FT DISULFID 304 325
FT CARBOHYD 56 56
FT VARSP LIC 34 88
FT AENPRLLEAEQKVFSSRGNGVTLPCFYRDPFAFGSGI
FT HKIRIKWLTSDY -> DCTAFWKILRGQRSSAPVGIL
FT TWPCCFPWRKHTYWKIGIKLPLSLAISDRTS (in
FT isoform Short).
FT /FTId=VSP 005301.
FT Missing (In isoform Short).
FT /FTId=VSP 005302.
FT R -> W (IN REF. 3).
FT VARSP LIC 89 158
FT CONFLICT 322 322
FT SEQUENCE 354 AA; 40261 MW; 23278A56273D6C CRC64;
SQ

```

```

Query Match 9.1%; Score 169.5; DB 1; Length 354;
Best Local Similarity 30.8%; Pred. No. 1.5e-07;
Matches 45; Conservative 20; Mismatches 52; Indels 29; Gaps 4;
QY 5 GKHKCEKSHVYVGDGLNCEPEQLPIDRCLQDNGQCHADAKVCLHFDQTTGVGVHLRSL 64
DB 135 GRYKE-----VIEGL-----EDTAVVALELQGV---VFYFPL 167
QY 65 GOYKLTDFKAREACANEATWATYNQLSYQKAKYHLCSAGWLETGRVAYPTAFASQNG 124
DB 168 GRYNLFHEARQACLDQDAVIASFDLYDAWGLGDLWCNAGWLSGVSQYPTTKPRECG 227
QY 125 --SGVVGIVDYGPRNKSEMDVFCY 148
DB 228 GQNTVPGVGRNYGFWDKSRDYDVCFC 253

RESULT 9
PLK_MOUSE
ID PLK_MOUSE STANDARD; PRT; 356 AA.
AC Q9QUP5; Q9D1G9; Q9Z1X7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Proteoglycan link protein precursor (Cartilage link protein) (LP).
GN CRTLL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1] _TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=20108790; PubMed=10640815;
RA Deak F., Mates L., Krysan K., Liu Z., Szabo P.E., Mann J.R.,
RA Beier D.R., Kiss I.;
RA "Characterization and chromosomal location of the mouse link protein
RT gene (Crtll).";
RL Cytogenet. Cell Genet. 87:75-79 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Otto J.M., Cs-Szabo G., Kamath R.V., Liu W., Li Y., Gkant T.T.;
RT "Molecular analysis of the murine link protein gene: 5' flanking,
RT coding and 3' flanking sequence analysis; exon-intron structure;
RT tissue distribution; and generation of transgenic mice.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085860; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -!- FUNCTION: Stabilizes the aggregates of proteoglycan monomers with
CC hyaluronic acid in the extracellular cartilage matrix.

```

RESULT 10

```

RESULT 11
ID BRAL HUMAN STANDARD; PRT; 340 AA.
AC Q9GZV7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Brain link protein-1 precursor.
GN BRAL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=20483618; PubMed=11027579;
RA Hirakawa S., Ohashi T., Su W.-D., Yoshioka H., Murakami T., Arata J.,
RA Nimomiya Y.;
RT "The brain link protein-1 (BRAL1): cDNA cloning, genomic structure,
RT and characterization as a novel link protein expressed in adult
RT brain.";
RL Biochem. Biophys. Res. Commun. 276:982-989 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Mediates a firm binding of versican V2 to hyaluronan-
CC acid. May play a pivotal role in the formation of the hyaluronan-
CC associated matrix in the central nervous system (CNS) which
CC facilitates neuronal conduction and general structural
CC stabilization. Binds to hyaluronan acid (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- TISSUE SPECIFICITY: Expressed only in adult brain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB049061; BAB17669.1;
CC EMBL; AB049057; BAB17669.1; JOINED.
CC EMBL; AB049058; BAB17669.1; JOINED.
CC EMBL; AB049059; BAB17669.1; JOINED.
CC EMBL; AB049060; BAB17669.1; JOINED.
CC EMBL; AB049054; BAB17662.1;

```

```

DR EMBL; BC029864; AAH29864.1; -.
DR PIR; JC7505; JC7505.
DR HSP; P98066; ITSG.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR Pfam; PF00047; Ig 1.
DR Pfam; PF00193; Xlink 2.
DR PRINTS; PR01285; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Proteoglycan; Hyaluronic acid; Extracellular matrix; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1..26 POTENTIAL.
FT CHAIN 27..340 BRAIN LINK PROTEIN-1.
FT DOMAIN 34..144 IG-LIKE V-TYPE.
FT DOMAIN 165..241 LINK 1.
FT DOMAIN 260..337 LINK 2.
FT DISULFID 57..128 BY SIMILARITY.
FT DISULFID 170..240 BY SIMILARITY.
FT DISULFID 194..215 BY SIMILARITY.
FT DISULFID 265..336 BY SIMILARITY.
FT DISULFID 290..311 BY SIMILARITY.
SQ SEQUENCE 340 AA; 37775 MW; 62CCCD86385F5C7 CRC64;

Query Match 8.9%; Score 166; DB 1; Length 340;
Best Local Similarity 31.7%; Pred. No. 2.9e-07;
Matches 40; Conservative 21; Mismatches 51; Indels 14; Gaps 4;

QY 33 LODNQCCHADAKCVDLH-FQDTTVG-----VFHLRSLPGQYKLTDFDKAREACANEA 83
DB 120 LEDEGR---YRCELINGIEDSEVALTILSLGFWPPYQPSRGYQFNYYEAKQACEEQDG 175
QY 84 TMTATYNQLSYQKAKYHLCSAGWLETGVAVTAFASQNGS-GVVGIVDYGPENKSEM 142
DB 176 RLATYSQLYQAWTEGLDNCWAGLLEGSVRYPVLTARAPCGGRGPRGIRSYGPRDRMRDR 235
QY 143 MDVFCY 148
DB 236 YDAFCF 241

RESULT 12
PLK_BOVIN STANDARD; PRT; 354 AA.
AC P55252;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Proteoglycan link protein precursor (Cartilage link protein) (LP).
GN CRT11.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=96043243; PubMed=7584851;
RA Hering T.M., Kollar J., Huynh T.D., Sandell L.J.;
RT "Bovine chondrocyte link protein cDNA sequence: interspecies
RT conservation of primary structure and mRNA untranslated regions.";
RL Comp. Biochem. Physiol. 112B:197-203 (1995).
CC -!- FUNCTION: Stabilizes the aggregates of proteoglycan monomers with
CC hyaluronic acid in the extracellular cartilage matrix.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02292; AAC04311.1; -.
CC HSSP; P98066; ITSG.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003596; IG v.
CC InterPro; IPR00538; Link.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00406; IGV; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS50835; IG-LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
CC Glycoprotein; Connective tissue; Extracellular matrix;
CC Hyaluronic acid; Proteoglycan; Immunoglobulin domain; Repeat.
FT PROPEP 1 15
FT CHAIN 16 354
FT DOMAIN 38 152
FT DOMAIN 176 253
FT DOMAIN 274 350
FT DISULFID 61 139
FT DISULFID 181 252
FT DISULFID 205 226
FT DISULFID 279 349
FT DISULFID 304 325
FT SEQUENCE 354 AA; 40287 MW; 06SD155378A1283C CRC64;
SQ
Query Match 8.9%; Score 165.5; DB 1; Length 354;
Best Local Similarity 30.1%; Pred. No. 3.4e-07;
Matches 44; Conservative 22; Mismatches 51; Indels 29; Gaps 4;
QY 5 GKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHLRSP 64
Db 135 GRYKCE-----VIEGL-----EDDTAVVALDLQGV---VFFYFRL 167
QY 65 GYKLTDFDKARACANEATMATYNQLSYXQKAYHLCAGWLETGRVAYPTAFASQNC 124
Db 168 GRYNLNFEHQAQCLDQDSTIASFDQLYEARSGLDWCNAGWLSGDSGVQYPTKPREPC 227
QY 125 --SGVGVIVDYGPRPNKSEMDVFCY 148
Db 228 GQNTVPGVRNYGFWDKSRDYDFCF 253
RESULT 13
ID PLK_CHICK STANDARD; PRT; 355 AA.
AC P07354;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Proteoglycan link protein precursor (Cartilage link protein) (LP).
GN CRT11.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryonic sternal cartilage;
RA MEDLINE=8623315; PubMed=3459154;
RA Deak F., Kiss I., Sparks K.J., Argraves W.S., Hampikian G.,
RA Goetinck P.F.;
RT "Complete amino acid sequence of chicken cartilage link protein
RT deduced from cDNA clones."

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 83:3766-3770(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87317659; PubMed=3476955;
RA Kiss I., Deak F., Mestric S., Delius H., Soos J., Dekany K.,
RA Argraves W.S., Sparks K.J., Goetinck P.F.;
RT "Structure of the chicken link protein gene: exons correlate with the
RT protein domains";
RL Proc. Natl. Acad. Sci. U.S.A. 84:6399-6403(1987).
CC -!- FUNCTION: Stabilizes the aggregates of proteoglycan monomers with
CC hyaluronic acid in the extracellular cartilage matrix.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M35038; AAA48941.1; -.
CC EMBL; M35035; AAA48941.1; JOINED.
CC EMBL; M35036; AAA48941.1; JOINED.
CC EMBL; M35037; AAA48941.1; JOINED.
CC EMBL; M13212; AAA48940.1; -.
CC PIR; A28305; LKCH.
CC HSSP; P98066; ITSG.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003596; IG v.
CC InterPro; IPR00538; Link.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00406; IGV; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS50835; IG-LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
CC Glycoprotein; Connective tissue; Extracellular matrix;
CC Hyaluronic acid; Proteoglycan; Immunoglobulin domain; Repeat.
FT PROPEP 1 9
FT CHAIN 10 355
FT DOMAIN 38 156
FT DOMAIN 177 254
FT DOMAIN 275 351
FT DISULFID 61 140
FT DISULFID 182 253
FT DISULFID 206 227
FT DISULFID 280 350
FT DISULFID 305 326
FT CARBOHYD 21 21
FT CARBOHYD 56 56
FT SEQUENCE 355 AA; 40533 MW; 2581CAE22158B60F CRC64;
SQ
Query Match 8.9%; Score 165.5; DB 1; Length 355;
Best Local Similarity 29.9%; Pred. No. 3.4e-07;
Matches 44; Conservative 26; Mismatches 46; Indels 31; Gaps 5;
QY 5 GKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHLRSP- 63
Db 136 GRYKCE-----VIEGL-----EDDTAVVALNLE---GVVFPYSR 167
QY 64 LGQYKLTDFDKARACANEATMATYNQLSYXQKAYHLCAGWLETGRVAYPTAFASQNC 123
Db 168 LGRYNLNFEHQAQCLDQDSTIASFDQLYEARSGLDWCNAGWLSGDSGVQYPTKPREPC 227
QY 124 G--SGVGVIVDYGPRPNKSEMDVFCY 148
Db 228 GQNTVPGVRNYGFWDKSRDYDFCF 254

```



```

CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB049055; BAB17663.1; -.
DR HSP; P98066; ITSG.
DR MGD; MGI:2137300; Brall.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR InterPro; IPR00538; Link.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Proteoglycan: Hyaluronic acid; Extracellular matrix; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 27
FT CHAIN 28 341
FT DOMAIN 35 143
FT DOMAIN 166 242
FT DOMAIN 261 338
FT DISULFID 58 129
FT DISULFID 171 241
FT DISULFID 195 216
FT DISULFID 266 337
FT DISULFID 291 312
SQ SEQUENCE 341 AA; 37925 MW; 1F65DCD07B5CFF02 CRC64;

Query Match 8.8%; Score 164; DB 1; Length 341;
Best Local Similarity 31.7%; Pred.No.4.4e-07;
Matches 40; Conservative 19; Mismatches 53; Indels 14; Gaps 4;

Qy 33 LQNGOCHADAKVCDLH-FQDTTVG-----VFHLRSPLGQYKLTDFDKAREACANEA 83
Db 121 LEDEGR-----YRCELNGIEDSEVALTLLEGVVFYQPSRGYQFNYEAKRACEQDG 176
Qy 84 TMATYNQLSYXKAKYHLCASGWLETRVAYPTAFASQNGS-GVVGIYDYGPRPNKSEM 142
Db 177 RLATYGQLYQAWTEGLDCWAGLLEGSVRYPLVTARPCGGHGRPGIRSYGPRDRSRD 236
Qy 143 WDVFCY 148
Db 237 YDAFCF 242

Search completed: March 30, 2004, 15:40:44
Job time : 18 secs

```

search completed:
Job time : 18 secs

1	262	14.0	277	1	US-08-024-868-2	Sequence 2, Appli
2	262	14.0	277	2	US-08-342-097-2	Sequence 2, Appli
3	262	14.0	277	3	US-09-206-695-2	Sequence 2, Appli
4	262	14.0	277	4	US-09-000-179-1	Sequence 1, Appli
5	262	14.0	277	4	US-09-799-118-2	Sequence 2, Appli
6	262	14.0	277	5	PCT-US96-11995-1	Sequence 1, Appli
7	187.5	10.1	360	4	US-09-907-7994A-213	Sequence 213, App
8	187.5	10.1	360	4	US-09-905-125A-213	Sequence 213, App
9	187.5	10.1	360	4	US-09-902-775A-213	Sequence 213, App
10	167	9.0	326	1	US-08-225-477B-6	Sequence 6, Appli
11	167	9.0	326	5	PCT-US95-04353-6	Sequence 6, Appli
12	160	8.6	98	2	US-08-242-097-4	Sequence 4, Appli
13	160	8.6	98	3	US-09-206-695-4	Sequence 4, Appli
14	160	8.6	98	4	US-09-799-118-4	Sequence 4, Appli
15	157.5	8.4	371	1	US-08-225-477B-8	Sequence 8, Appli
16	157.5	8.4	371	5	PCT-US95-04353-8	Sequence 8, Appli
17	157.5	8.4	1257	1	US-08-340-428B-49	Sequence 49, Appli
18	157	8.4	329	1	US-08-225-477B-3	Sequence 3, Appli
19	157	8.4	329	5	PCT-US95-04353-3	Sequence 3, Appli
20	156	8.4	333	1	US-08-225-477B-4	Sequence 4, Appli
21	156	8.4	333	5	PCT-US95-04353-4	Sequence 4, Appli
22	155.5	8.3	912	5	PCT-US95-03747-2	Sequence 2, Appli
23	155	8.3	528	4	US-09-010-147B-20	Sequence 20, Appli
24	153.5	8.2	97	2	US-08-242-097-5	Sequence 5, Appli
25	153.5	8.2	97	3	US-09-206-695-5	Sequence 5, Appli
26	153.5	8.2	97	4	US-09-799-118-5	Sequence 5, Appli
27	150.5	8.1	328	1	US-08-225-477B-5	Sequence 5, Appli

QY 52 DTTVGVFHLRSP LGQYKLT FDKAREACANEAA TMATYNQLSYXQKAKYHLC SAGWLETGR 111

Db 32 ERAAGVYHREARSGKYLTAEAKAVCEFFEGGHLATYKQLEAARKIGFHVCAAGWMAKGR 91
QY 112 VAYPTAFASQNGSGVGVIGVYDYGPRPNKSEMDVFCYRMKDVC 155
Db 92 VGYPIVKGPCNGCGKGTGIIDYGIIRNRSERWDAYCYNPHAKEC 135

RESULT 2

US-08-242-097-2
; Sequence 2, Application US/08242097
; Patent No. 5846763
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; TITLE OF INVENTION: Therefor and Uses Thereof
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; TELEFAX: 212-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-242-097-2

Query Match 14.0%; Score 262; DB 2; Length 277;
Best Local Similarity 43.3%; Pred. No. 4.1e-20;
Matches 45; Conservative 16; Mismatches 43; Indels 0; Gaps 0;
QY 52 DTTGVFHLRPLGQYKLTFDKAREACANEATMATYNQLSYXOKAKYHLCASAGWLETGR 111
Db 32 ERAAGVYHREARSGKYLTAEAKAVCEFFEGGHLATYKQLEAARKIGFHVCAAGWMAKGR 91
QY 112 VAYPTAFASQNGSGVGVIGVYDYGPRPNKSEMDVFCYRMKDVC 155
Db 92 VGYPIVKGPCNGCGKGTGIIDYGIIRNRSERWDAYCYNPHAKEC 135

RESULT 3

US-09-206-695-2
; Sequence 2, Application US/09206695
; Patent No. 6210905
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; TITLE OF INVENTION: Therefor and Uses Thereof
; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; TELEFAX: 212-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-206-695-2

Query Match 14.0%; Score 262; DB 3; Length 277;
Best Local Similarity 43.3%; Pred. No. 4.1e-20;
Matches 45; Conservative 16; Mismatches 43; Indels 0; Gaps 0;
QY 52 DTTGVFHLRPLGQYKLTFDKAREACANEATMATYNQLSYXOKAKYHLCASAGWLETGR 111
Db 32 ERAAGVYHREARSGKYLTAEAKAVCEFFEGGHLATYKQLEAARKIGFHVCAAGWMAKGR 91
QY 112 VAYPTAFASQNGSGVGVIGVYDYGPRPNKSEMDVFCYRMKDVC 155
Db 92 VGYPIVKGPCNGCGKGTGIIDYGIIRNRSERWDAYCYNPHAKEC 135

RESULT 4

US-09-000-179-1
; Sequence 1, Application US/09000179
; Patent No. 6313091
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Pharmaceutical Composition
; TITLE OF INVENTION: Containing TSG-6 Protein for Treating Inflammatory Diseases and
; TITLE OF INVENTION: Cancer-Related Pathologies and Methods of Using same
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,179
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,311

RESULT 7

GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 213
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-794A-213

Query Match 10.1%; Score 187.5; DB 4; Length 360;

Best Local Similarity 30.8%; Pred. No. 8.1e-12;
Matches 45; Conservative 24; Mismatches 48; Indels 29; Gaps 5;
QY 5 GKHKCECKSHVYGDGLNCEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTGVFHLRSL 64
Db 142 GRYRCE-----VIDGLEDE-----SGLVELELRGV---VFPYQSPN 174
QY 65 GQYKLTFTKAREACANEATMATYNQLSYXQKAKYHLCASAGWLETGRVAYETAFASQNCG 124
Db 175 GRYQFNFEHQGVCAEQAAVVASPEQLFRWEEGLDWCNAGWLQDATVQYPIMLPRQPCG 234
QY 125 S-GTV-GIVDYGPRPNKSEMMWDFCY 148
Db 235 GPGLAPGVRSYGPRRRLHRYDVFECF 260

RESULT 8
US-09-905-125A-213
; Sequence 213, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30

```
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 213
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-125A-213

Query Match      10.1%; Score 187.5; DB 4; Length 360;
Best Local Similarity 30.8%; Pred. No. 8.1e-12;
Matches 45; Conservative 24; Mismatches 48; Indels 29; Gaps 5;

Qy      5 GKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKVDLHFQDTTGVFHLRSP 64
Db      142 GRYRCE-----VIDGLEDE-----SGLVELELRGV---VFPYQSPN 174

Qy      65 GOYKLTFDKAREACANEATWATYNQLSYOKAKYHLCAGWLETGRVAYPTAFASQNG 124
Db      175 GRYQPNFHEGQQVCAEQAAVVASPEQLFRAWEGLDWCNAGWLQDATVQYPIMLPRQPCG 234

Qy      125 S-GVV-GIVDYGPRPNKSEMDVFCY 148
Db      235 GPGLAPGVRSYGPRHRLHRYDVFCF 260

RESULT 9
US-09-902-775A-213
; Sequence 213, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavir, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Pao, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
```

```
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 213
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-775A-213

Query Match      10.1%; Score 187.5; DB 4; Length 360;
Best Local Similarity 30.8%; Pred. No. 8.1e-12;
Matches 45; Conservative 24; Mismatches 48; Indels 29; Gaps 5;

Qy      5 GKHKCECKSHVVGDLNCEPEQLPIDRCLQDNGQCHADAKVDLHFQDTTGVFHLRSP 64
Db      142 GRYRCE-----VIDGLEDE-----SGLVELELRGV---VFPYQSPN 174

Qy      65 GOYKLTFDKAREACANEATWATYNQLSYOKAKYHLCAGWLETGRVAYPTAFASQNG 124
Db      175 GRYQPNFHEGQQVCAEQAAVVASPEQLFRAWEGLDWCNAGWLQDATVQYPIMLPRQPCG 234

Qy      125 S-GVV-GIVDYGPRPNKSEMDVFCY 148
Db      235 GPGLAPGVRSYGPRHRLHRYDVFCF 260

RESULT 10
US-08-225-477B-6
; Sequence 6, Application US/08225477B
; Patent No. 5635370
; GENERAL INFORMATION:
; APPLICANT: Susan Hockfield
; APPLICANT: Diane M. Jaworski
; TITLE OF INVENTION: BEHAB, A Brain Hyal
; TITLE OF INVENTION: luronan-Binding Protein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: St. Onge Steward Johnston & Reens
; STREET: 986 Bedford Street
; CITY: Stamford
; STATE: CT
; COUNTRY: United States
; ZIP: 06905
```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/225,477B
FILING DATE: April 8, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: 1751-P0004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-324-6155
TELEFAX: 203-327-1096
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: polypeptide
FRAGMENT TYPE: functional domains
FEATURE:
NAME/KEY: rat link protein
PUBLICATION INFORMATION:
AUTHORS: Doege, K., Hassell, J.R., Ca-
AUTHORS: terson, B., and Yamada, Y.
TITLE: Link protein cDNA sequence reveals a
TITLE: tandemly repeated protein sequence.
JOURNAL: Proc. Natl. Acad. Sci. USA
VOLUME: 83
PAGES: 3761-3765
DATE: 1986
US-08-225-477B-6

Query Match 9.0%; Score 167; DB 1; Length 326;

Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 46; Conservative 21; Mismatches 49; Indels 30; Gaps 5;

QY 5 GKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGHLSPL 64
Db 109 GRYKCE-----VIEGL-----EDTAVVALELQGV---VFPYFREL 141
QY 65 GQYKLTDFKARACANEATMTYNQLSYQKAKYHLCAGMLETGRVAYPTAFASQNG 124
Db 142 GRYNLFHEARQACLDQDAVIASFDQLYDAWRGGLDWCNAGWLSGDSGVQYPTTKPREPCG 201
QY 125 --SGVVGVVDYGPRLNKSEMDVFCY 148
Db 202 GQNTVPGVGRNYG-FWDKDSRYDVFCF 226

RESULT 11

PCT-US95-04353-6
Sequence 6, Application PC/TUS9504353
GENERAL INFORMATION:
APPLICANT: Susan Hockfield
APPLICANT: Diane M. Jaworski
TITLE OF INVENTION: BEHAB, A Brain Hya-
TITLE OF INVENTION: luteonin-Binding Protein
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: St. Onge Steward Johnston & Reens
STREET: 986 Bedford Street
CITY: Stamford
STATE: CT
COUNTRY: United States
ZIP: 06905
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC

OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04353
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,477
FILING DATE: April 8, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: 1751-P0004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-324-6155
TELEFAX: 203-327-1096
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: polypeptide
FRAGMENT TYPE: functional domains
FEATURE:
NAME/KEY: rat link protein
PUBLICATION INFORMATION:
AUTHORS: Doege, K., Hassell, J.R., Ca-
AUTHORS: terson, B., and Yamada, Y.
TITLE: Link protein cDNA sequence reveals a
TITLE: tandemly repeated protein sequence.
JOURNAL: Proc. Natl. Acad. Sci. USA
VOLUME: 83
PAGES: 3761-3765
DATE: 1986
PCT-US95-04353-6

Query Match 9.0%; Score 167; DB 5; Length 326;

Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 46; Conservative 21; Mismatches 49; Indels 30; Gaps 5;

QY 5 GKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGHLSPL 64
Db 109 GRYKCE-----VIEGL-----EDTAVVALELQGV---VFPYFREL 141
QY 65 GQYKLTDFKARACANEATMTYNQLSYQKAKYHLCAGMLETGRVAYPTAFASQNG 124
Db 142 GRYNLFHEARQACLDQDAVIASFDQLYDAWRGGLDWCNAGWLSGDSGVQYPTTKPREPCG 201
QY 125 --SGVVGVVDYGPRLNKSEMDVFCY 148
Db 202 GQNTVPGVGRNYG-FWDKDSRYDVFCF 226

RESULT 12

US-08-242-097-4
Sequence 4, Application US/08242097
Patent No. 5846763
GENERAL INFORMATION:
APPLICANT: Lee, Tae Ho
APPLICANT: Wisniewski, Hans Georg
APPLICANT: Vilcek, Jan
TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
TITLE OF INVENTION: Therefor and Uses Thereof
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-242-097-4

Query Match      8.6%; Score 160; DB 2; Length 98;
Best Local Similarity 36.2%; Pred. No. 1.1e-09;
Matches 34; Conservative 17; Mismatches 41; Indels 2; Gaps 1;

QY 57 VFHLRSPLGQYKLTDFDKAREACANEATMATYNQLSYXOKAKYHLCSAGWLETGRVAYPT 116
Db 3 VFYPFRLGRYLNLPHEARQACLDQDAVIASFDQLYDAWRGLDWCNAGWLSDGSVQYPI 62

QY 117 AFASQNCG--SGVVGIVDYGPRPNKSEMMWDVFCY 148
Db 63 TKPREPCGGQNTVPGVRNYGFWDKDKSRYDVFCF 96

RESULT 13
US-09-206-695-4
; Sequence 4, Application US/09206695
; Patent No. 6210905
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-206-695-4

Query Match      8.6%; Score 160; DB 2; Length 98;
Best Local Similarity 36.2%; Pred. No. 1.1e-09;
Matches 34; Conservative 17; Mismatches 41; Indels 2; Gaps 1;

QY 57 VFHLRSPLGQYKLTDFDKAREACANEATMATYNQLSYXOKAKYHLCSAGWLETGRVAYPT 116
Db 3 VFYPFRLGRYLNLPHEARQACLDQDAVIASFDQLYDAWRGLDWCNAGWLSDGSVQYPI 62

QY 117 AFASQNCG--SGVVGIVDYGPRPNKSEMMWDVFCY 148
Db 63 TKPREPCGGQNTVPGVRNYGFWDKDKSRYDVFCF 96

RESULT 13
US-09-206-695-4
; Sequence 4, Application US/09206695
; Patent No. 6210905
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-206-695-4
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-206-695-4

Query Match      8.6%; Score 160; DB 3; Length 98;
Best Local Similarity 36.2%; Pred. No. 1.1e-09;
Matches 34; Conservative 17; Mismatches 41; Indels 2; Gaps 1;

QY 57 VFHLRSPLGQYKLTDFDKAREACANEATMATYNQLSYXOKAKYHLCSAGWLETGRVAYPT 116
Db 3 VFYPFRLGRYLNLPHEARQACLDQDAVIASFDQLYDAWRGLDWCNAGWLSDGSVQYPI 62

QY 117 AFASQNCG--SGVVGIVDYGPRPNKSEMMWDVFCY 148
Db 63 TKPREPCGGQNTVPGVRNYGFWDKDKSRYDVFCF 96

RESULT 14
US-09-799-118-4
; Sequence 4, Application US/09799118
; Patent No. 6518401
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,118
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-799-118-4

Query Match      8.6%; Score 160; DB 4; Length 98;
Best Local Similarity 36.2%; Pred. No. 1.1e-09;
Matches 34; Conservative 17; Mismatches 41; Indels 2; Gaps 1;

QY 57 VFHLRSPLGQYKLTDFDKAREACANEATMATYNQLSYXOKAKYHLCSAGWLETGRVAYPT 116
Db 3 VFYPFRLGRYLNLPHEARQACLDQDAVIASFDQLYDAWRGLDWCNAGWLSDGSVQYPI 62
```


No art
pink words
Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 30, 2004, 15:42:36 ; Search time 45 Seconds
(without alignments)
2052.591 Million cell updates/sec

Title: US-09-466-778B-11
Perfect score: 1865
Sequence: 1 MTGPKHKCEKSHVGDGL.....ALAAYSYPRINKTIGFXHP 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1772	95.0	1192	15	US-10-028-248A-40
2	1772	95.0	1192	15	US-10-107-782-40
3	1772	95.0	1394	9	US-09-842-930A-25
4	1772	95.0	1416	14	US-10-133-172-4
5	1772	95.0	1653	14	US-10-133-172-20
6	1769	94.9	897	15	US-10-028-248A-41
7	1769	94.9	897	15	US-10-028-248A-211
8	1769	94.9	897	15	US-10-107-782-41
9	1769	94.9	897	15	US-10-107-782-211
10	1554	83.3	2675	15	US-10-028-248A-2
11	1554	83.3	2675	15	US-10-107-782-2
12	1514.5	81.2	1431	9	US-09-842-930A-2
13	1514.5	81.2	1431	14	US-10-133-172-2
14	1480	79.4	1069	15	US-10-028-248A-39
15	1480	79.4	1069	15	US-10-107-782-39

16	847	45.4	2420	15	US-10-028-248A-4
17	847	45.4	2420	15	US-10-107-782-4
18	668	35.8	2212	15	US-10-028-248A-43
19	668	35.8	2212	15	US-10-107-782-43
20	661.5	35.5	1082	15	US-10-264-237-2725
21	652.5	35.0	2570	15	US-10-028-248A-42
22	652.5	35.0	2570	15	US-10-107-782-42
23	645.5	34.6	457	10	US-09-774-639-108
24	645.5	34.6	457	10	US-09-969-730-110
25	645.5	34.6	457	16	US-10-621-363-110
26	642	34.4	482	15	US-10-104-047-3794
27	581	31.2	510	15	US-10-104-047-2580
28	505	27.1	93	14	US-10-133-172-5
29	262	14.0	277	9	US-09-799-118-2
30	262	14.0	277	14	US-10-247-671-185
31	262	14.0	277	15	US-10-295-027-292
32	262	14.0	277	15	US-10-295-027-1237
33	244	13.1	277	15	US-10-295-027-294
34	213	11.4	106	14	US-10-023-896-82
35	196	10.5	104	15	US-10-028-248A-46
36	196	10.5	104	15	US-10-107-782-46
37	187.5	10.1	360	9	US-09-909-320-213
38	187.5	10.1	360	9	US-09-909-088B-213
39	187.5	10.1	360	9	US-09-905-291A-213
40	187.5	10.1	360	9	US-09-902-853-213
41	187.5	10.1	360	9	US-09-907-824-213
42	187.5	10.1	360	9	US-09-907-841-213
43	187.5	10.1	360	10	US-09-904-011-213
44	187.5	10.1	360	10	US-09-906-742-213
45	187.5	10.1	360	10	US-09-906-838-213

ALIGNMENTS

RESULT 1
US-10-028-248A-40
; Sequence 40, Application US/10028248A
; Publication NO. US20030235882A1
; GENERAL INFORMATION:

- APPLICANT: Shimkets, Richard
- APPLICANT: Patturajan, Meera
- APPLICANT: Vernet, Corine
- APPLICANT: Casman, Stacie
- APPLICANT: Malyankar, Uriel
- APPLICANT: Shenoy, Suresh
- APPLICANT: Spytek, Kimberly
- APPLICANT: Gangolli, Esha
- APPLICANT: Miller, Charles
- APPLICANT: Boldog, Ferenc
- APPLICANT: Li, Li
- APPLICANT: Taupier Jr, Raymond J
- APPLICANT: Kekuda, Ramesh
- APPLICANT: Smithson, Glennnda
- APPLICANT: Zerhusen, Bryan
- APPLICANT: Liu, Xiaohong
- APPLICANT: Colman, Steven
- APPLICANT: Tchernev, Velizar
- APPLICANT: Si, Jingsheng
- APPLICANT: Edinger, Shlomit
- APPLICANT: Stone, David
- APPLICANT: Sciore, Paul
- APPLICANT: Millet, Isabelle
- APPLICANT: Rothenberg, Mark
- TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
- FILE OF INVENTION: 21402-222
- CURRENT APPLICATION NUMBER: US/10/028,248A
- CURRENT FILING DATE: 2001-12-19
- PRIOR APPLICATION NUMBER: 60/256619
- PRIOR FILING DATE: 2000-12-19
- PRIOR APPLICATION NUMBER: 60/262959
- PRIOR FILING DATE: 2001-01-19

```
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-40

Query Match          95.0%; Score 1772; DB 15; Length 1192;
Best Local Similarity 94.6%; Pred. No. 1.5e-183;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 784 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 843

QY 61 RSPGQYKLTDFKAREACANEATMATYNQLSYXQAKYHLCSAGWLETGRVAYPTAFAS 120
Db 844 RSPGQYKLTDFKAREACANEATMATYNQLSYXQAKYHLCSAGWLETGRVAYPTAFAS 903

QY 121 QNCGSGVGIYDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFSYGNLLQVLMGFP 180
Db 904 QNCGSGVGIYDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFSYGNLLQVLMGFP 963

QY 181 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
Db 964 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1023

QY 241 VSMFFYNDLVNGTTLQTRLGSKLLITDRQDPLHPTETRCVDRDLEWDICASNGITHVI 300
Db 1024 VSMFFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIIHVI 1083

QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353
Db 1084 SRXLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSFRINRRTTIGFQHF 1136

RESULT 2
US-10-107-782-40
; Sequence 40, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,

; APPLICANT: Zethusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 40
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-40

Query Match          95.0%; Score 1772; DB 15; Length 1192;
Best Local Similarity 94.6%; Pred. No. 1.5e-183;
Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 784 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 843

QY 61 RSPGQYKLTDFKAREACANEATMATYNQLSYXQAKYHLCSAGWLETGRVAYPTAFAS 120
Db 844 RSPGQYKLTDFKAREACANEATMATYNQLSYXQAKYHLCSAGWLETGRVAYPTAFAS 903

QY 121 QNCGSGVGIYDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFSYGNLLQVLMGFP 180
Db 904 QNCGSGVGIYDYGPRPNKSEMDVFCYRMKDVNCTCKVGYVGDGFSYGNLLQVLMGFP 963

QY 181 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
Db 964 SLTNFLTEVLAYSNSARGRAFLHLEHTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1023

QY 241 VSMFFYNDLVNGTTLQTRLGSKLLITDRQDPLHPTETRCVDRDLEWDICASNGITHVI 300
Db 1024 VSMFFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIIHVI 1083

QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353
Db 1084 SRXLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSFRINRRTTIGFQHF 1136

RESULT 3
US-09-842-930A-25
; Sequence 25, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
```

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 1394

; TYPE: PR

; ORGANISM: Homo sapiens

US-09-842-930A-25

Query Match

; Sequence 4, Application US/10133172; DB 9; Length 1394;

; Best Local Similarity 94.6%; Pred. No. 1.9e-183;

Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 60

Db 986 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 1045

Qy 61 RSPGQYKLTDFDKAREACANEAAATWATYNQLSYXQAKYHLCISAGWLETGRVAYPTAFAS 120

Db 1046 RSPGQYKLTDFDKAREACANEAAATWATYNQLSYXQAKYHLCISAGWLETGRVAYPTAFAS 1105

Qy 121 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVGDGFSYSGNLLQVLMSPF 180

Db 1106 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVGDGFSYSGNLLQVLMSPF 1165

Qy 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTGLFVPQNSGLGENETLSGRDIEHHLAN 240

Db 1166 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTGLFVPQNSGLGENETLSGRDIEHHLAN 1225

Qy 241 VSMFFYNDLVNGTTLTQRLGSKLLITDRQDPLHPTETRCVDGRDTLEWDICASNGIITHVI 300

Db 1226 VSMFFYNDLVNGTTLTQRLGSKLLITASQDPLQPTETRFVDGRALQWDIFASNGIITHVI 1285

Qy 301 SRXLKAPPAPVTLKHTGLGXGIFXXIIILVTGAVALAAYSYFRINRKTIGFQHF 353

Db 1286 SRPLKAPPAPVTLKHTGLGXGIFFAIILVTGAVALAAYSYFRINRRTIGFQHF 1338

RESULT 4

US-10-133-172-4

; Sequence 4, Application US/10133172

; Publication No. US20030104987A1

; GENERAL INFORMATION:

; APPLICANT: WEIGEL, PAUL H

; APPLICANT: WEIGEL, JANET A

; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS

; FILE REFERENCE: 5864.014

; CURRENT APPLICATION NUMBER: US/10/133,172

; PRIOR FILING DATE: 2002-04-25

; PRIOR FILING DATE: 2001-04-25

; PRIOR FILING DATE: 2001-04-25

; PRIOR FILING DATE: 2001-04-25

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1416

; TYPE: PR

; ORGANISM: Homo sapiens

US-10-133-172-4

Query Match

; Sequence 4, Application US/10133172; DB 14; Length 1416;

; Best Local Similarity 94.6%; Pred. No. 2e-183;

Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 60

Db 1008 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 1067

Qy 61 RSPGQYKLTDFDKAREACANEAAATWATYNQLSYXQAKYHLCISAGWLETGRVAYPTAFAS 120

Db 1068 RSPGQYKLTDFDKAREACANEAAATWATYNQLSYXQAKYHLCISAGWLETGRVAYPTAFAS 1127

Qy 121 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVGDGFSYSGNLLQVLMSPF 180

Db 1286 SRPLKAPPAPVTLKHTGLGXGIFFAIILVTGAVALAAYSYFRINRRTIGFQHF 1338

Db 1128 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVGDGFSYSGNLLQVLMSPF 1187

Qy 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTGLFVPQNSGLGENETLSGRDIEHHLAN 240

Db 1188 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTGLFVPQNSGLGENETLSGRDIEHHLAN 1247

Qy 241 VSMFFYNDLVNGTTLTQRLGSKLLITDRQDPLHPTETRCVDGRDTLEWDICASNGIITHVI 300

Db 1248 VSMFFYNDLVNGTTLTQRLGSKLLITASQDPLQPTETRFVDGRALQWDIFASNGIITHVI 1307

Qy 301 SRXLKAPPAPVTLKHTGLGXGIFXXIIILVTGAVALAAYSYFRINRKTIGFQHF 353

Db 1308 SRPLKAPPAPVTLKHTGLGXGIFFAIILVTGAVALAAYSYFRINRRTIGFQHF 1360

RESULT 5

US-10-133-172-20

; Sequence 20, Application US/10133172

; Publication No. US20030104987A1

; GENERAL INFORMATION:

; APPLICANT: WEIGEL, PAUL H

; APPLICANT: WEIGEL, JANET A

; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS

; FILE REFERENCE: 5864.014

; CURRENT APPLICATION NUMBER: US/10/133,172

; PRIOR FILING DATE: 2002-04-25

; PRIOR FILING DATE: 2001-04-25

; PRIOR FILING DATE: 2001-04-25

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 20

; LENGTH: 1653

; TYPE: PR

; ORGANISM: Homo sapiens

US-10-133-172-20

Query Match

; Sequence 20, Application US/10133172; DB 14; Length 1653;

; Best Local Similarity 94.6%; Pred. No. 2.5e-183;

Matches 334; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 60

Db 1245 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 1304

Qy 61 RSPGQYKLTDFDKAREACANEAAATWATYNQLSYXQAKYHLCISAGWLETGRVAYPTAFAS 120

Db 1305 RSPGQYKLTDFDKAREACANEAAATWATYNQLSYXQAKYHLCISAGWLETGRVAYPTAFAS 1364

Qy 121 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVGDGFSYSGNLLQVLMSPF 180

Db 1365 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTCKVGVGDGFSYSGNLLQVLMSPF 1424

Qy 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTGLFVPQNSGLGENETLSGRDIEHHLAN 240

Db 1425 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTGLFVPQNSGLGENETLSGRDIEHHLAN 1484

Qy 241 VSMFFYNDLVNGTTLTQRLGSKLLITDRQDPLHPTETRCVDGRDTLEWDICASNGIITHVI 300

Db 1485 VSMFFYNDLVNGTTLTQRLGSKLLITASQDPLQPTETRFVDGRALQWDIFASNGIITHVI 1544

Qy 301 SRXLKAPPAPVTLKHTGLGXGIFXXIIILVTGAVALAAYSYFRINRKTIGFQHF 353

Db 1545 SRPLKAPPAPVTLKHTGLGXGIFFAIILVTGAVALAAYSYFRINRRTIGFQHF 1597

RESULT 6

US-10-028-248A-41

; Sequence 41, Application US/10028248A

; Publication No. US20030235882A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard

APPLICANT: Patturajan, Meera
APPLICANT: Vernet, Corine
APPLICANT: Casman, Stacie
APPLICANT: Malyankar, Uriel
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Gangolli, Esha
APPLICANT: Miller, Charles
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Kekuda, Ramesh
APPLICANT: Smithson, Glennnda
APPLICANT: Zethusen, Bryan
APPLICANT: Liu, Xiaohong
APPLICANT: Colman, Steven
APPLICANT: Tchernev, Velizar
APPLICANT: Si, Jingsheng
APPLICANT: Edinger, Shlomit
APPLICANT: Stone, David
APPLICANT: Sciore, Paul
APPLICANT: Millet, Isabelle
APPLICANT: Rothenberg, Mark
TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
TITLE OF INVENTION: Thereof
FILE REFERENCE: 21402-222
CURRENT APPLICATION NUMBER: US/10/028,248A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/256619
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/262959
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/272408
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/285189
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: 60/308039
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: 60/311266
PRIOR FILING DATE: 2001-08-09
NUMBER OF SEQ ID NOS: 211
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 897
TYPE: PRT
ORGANISM: Homo sapiens
US-10-028-248A-41

Query Match 94.9%; Score 1769; DB 15; Length 897;
Best Local Similarity 94.3%; Pred. No. 2.1e-183;
Matches 333; Conservative 3; Mismatches 17; Indels 0; Gaps 0;
QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQNGQCHADAKCVDLHFQDITTVGVFHL 60
DB 489 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQNGQCHADAKCVDLHFQDITTVGVFHL 548
QY 61 RSPLOQYKLTFDKAREACANEATMATYNQLSYQAKYHLCASAGWLETGRVAYPTAFAS 120
DB 549 RSPLOQYKLTFDKAREACANEATMATYNQLSYQAKYHLCASAGWLETGRVAYPTAFAS 608
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDGFSYSGNLLQVLMSPF 180
DB 609 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDGFSYSGNLLQVLMSPF 668
QY 181 SLTNFLTEVLAYSNSARGRAFLEHLTDLISIRGTLFVQNSGLGENETLSGRDIEHHLAN 240
DB 669 SLTNFLTEVLAYSNSARGRAFLEHLTDLISIRGTLFVQNSGLGENETLSGRDIEHHLAN 728
QY 241 VSMFFNDLVNGTTLQTRGLSKLLITDRQDPLHPETRCVDDORDLEWDICASNGITHVI 300
DB 729 VSMFFNDLVNGTTLQTRGLSKLLITASQDPLQPTETREVDGRAILLQWDFASNGIIHVI 788
QY 301 SRXLKAPPAPVTLTHTGLGXGIFXXIILVTGAVALAAYSYPFRNKRTIGFXHF 353

DB 789 SRPLKAPPAPVTLTHTGLGAGIFFAILVTGAVALAAYSYPFRNKRTIGFQHF 841

RESULT 7

US-10-028-248A-211
Sequence 211, Application US/10028248A
Publication No. US20030235882A1

GENERAL INFORMATION:

APPLICANT: Shinkets, Richard
APPLICANT: Patturajan, Meera
APPLICANT: Vernet, Corine
APPLICANT: Casman, Stacie
APPLICANT: Malyankar, Uriel
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Gangolli, Esha
APPLICANT: Miller, Charles
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Kekuda, Ramesh
APPLICANT: Smithson, Glennnda
APPLICANT: Zethusen, Bryan
APPLICANT: Liu, Xiaohong
APPLICANT: Colman, Steven
APPLICANT: Tchernev, Velizar
APPLICANT: Si, Jingsheng
APPLICANT: Edinger, Shlomit
APPLICANT: Stone, David
APPLICANT: Sciore, Paul
APPLICANT: Millet, Isabelle
APPLICANT: Rothenberg, Mark
TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
TITLE OF INVENTION: Thereof
FILE REFERENCE: 21402-222
CURRENT APPLICATION NUMBER: US/10/028,248A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/256619
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/262959
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/272408
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/285189
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: 60/308039
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: 60/311266
PRIOR FILING DATE: 2001-08-09
NUMBER OF SEQ ID NOS: 211
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 211
LENGTH: 897
TYPE: PRT
ORGANISM: Homo sapiens
US-10-028-248A-211

Query Match 94.9%; Score 1769; DB 15; Length 897;
Best Local Similarity 94.3%; Pred. No. 2.1e-183;
Matches 333; Conservative 3; Mismatches 17; Indels 0; Gaps 0;
QY 1 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQNGQCHADAKCVDLHFQDITTVGVFHL 60
DB 489 MTGPGKHCKESHVYVGDGLNCEPEQLPIDRCLQNGQCHADAKCVDLHFQDITTVGVFHL 548
QY 61 RSPLOQYKLTFDKAREACANEATMATYNQLSYQAKYHLCASAGWLETGRVAYPTAFAS 120
DB 549 RSPLOQYKLTFDKAREACANEATMATYNQLSYQAKYHLCASAGWLETGRVAYPTAFAS 608
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDGFSYSGNLLQVLMSPF 180
DB 609 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTKXVGVGDGFSYSGNLLQVLMSPF 668

QY 181 SUTNFLTVEVLAAYSSSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 DB 669 SUTNFLTVEVLAAYSSSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 728
 QY 241 VSMFFYNDLVNGTTLTQTRVGSKLLITDRODPLHPTETRCVGDGRDTEWDICASNGITHVI 300
 DB 729 VSMFFYNDLVNGTTLTQTRVGSKLLITASQDPLQPTETRFVDGRAILOWDIFASNGIHHVI 788
 QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGVALAAYSYFRINRKTIGFXHF 353
 DB 789 SRPLKAPPAPVTLTHGLGAGIFFAILVTGVALAAYSYFRINRKTIGFQHF 841

RESULT 8

US-10-107-782-41
 ; Sequence 41, Application US/10107782
 ; Publication No. US20040018970A1

GENERAL INFORMATION:

; APPLICANT: Boldog, Ferenc,
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Colman, Steve,
 ; APPLICANT: Edinger, Shlomit,
 ; APPLICANT: Gangolli, Esha,
 ; APPLICANT: Kekuda, Ramesh,
 ; APPLICANT: Li, Li,
 ; APPLICANT: Liu, Xiaohong,
 ; APPLICANT: Malyankar, Uriel,
 ; APPLICANT: Miller, Charles,
 ; APPLICANT: Millet, Isabelle,
 ; APPLICANT: Patturajan, Meera,
 ; APPLICANT: Rothenberg, Mark,
 ; APPLICANT: Sciore, Paul,
 ; APPLICANT: Shenoy, Suresh,
 ; APPLICANT: Shimkets, Richard,
 ; APPLICANT: Si, Jingsheng,
 ; APPLICANT: Smithson, Glennda,
 ; APPLICANT: Spytek, Kimberly,
 ; APPLICANT: Stone, David,
 ; APPLICANT: Taupier, Raymond, jr.,
 ; APPLICANT: Tchernev, Velizar,
 ; APPLICANT: Vernet, Corine,
 ; APPLICANT: Zerhusen, Brian

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF

; FILE REFERENCE: 21402-222CIP

; CURRENT APPLICATION NUMBER: US/10/107,782

; PRIOR FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: 10/028,248

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/256,619

; PRIOR FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: 60/262,959

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/272,408

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 60/285,189

; PRIOR FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: 60/308,039

; PRIOR FILING DATE: 2001-07-26

; PRIOR APPLICATION NUMBER: 60/311,266

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/279,344

; PRIOR FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 215

; SOFTWARE: CuraSeqlist version 0.1

; SEQ ID NO 41

; LENGTH: 897

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-107-782-41

Query Match

Best Local Similarity 94.9%; Score 1769; DB 15; Length 897;

Pred. No. 2.1e-183;

Matches 333; Conservative 3; Mismatches 17; Indels 0; Gaps 0;
 QY 1 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
 DB 489 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 548
 QY 61 RSPLCQYKLTFTDKAREACANEAAATWATNQLSYXOKAKYHLCSAGWLETGRVAYPTAFAS 120
 DB 549 RSPLCQYKLTFTDKAREACANEAAATWATNQLSYXOKAKYHLCSAGWLETGRVAYPTAFAS 608
 QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVCNTXKVGVYVDGFSYSGNLLQVLMSPF 180
 DB 609 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVCNTXKVGVYVDGFSYSGNLLQVLMSPF 668
 QY 181 SUTNFLTVEVLAAYSSSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
 DB 669 SUTNFLTVEVLAAYSSSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 728
 QY 241 VSMFFYNDLVNGTTLTQTRVGSKLLITDRODPLHPTETRCVGDGRDTEWDICASNGITHVI 300
 DB 729 VSMFFYNDLVNGTTLTQTRVGSKLLITASQDPLQPTETRFVDGRAILOWDIFASNGIHHVI 788
 QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGVALAAYSYFRINRKTIGFXHF 353
 DB 789 SRPLKAPPAPVTLTHGLGAGIFFAILVTGVALAAYSYFRINRKTIGFQHF 841

RESULT 9

US-10-107-782-211

; Sequence 211, Application US/10107782

; Publication No. US20040018970A1

GENERAL INFORMATION:

; APPLICANT: Boldog, Ferenc,
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Colman, Steve,
 ; APPLICANT: Edinger, Shlomit,
 ; APPLICANT: Gangolli, Esha,
 ; APPLICANT: Kekuda, Ramesh,
 ; APPLICANT: Li, Li,
 ; APPLICANT: Liu, Xiaohong,
 ; APPLICANT: Malyankar, Uriel,
 ; APPLICANT: Miller, Charles,
 ; APPLICANT: Millet, Isabelle,
 ; APPLICANT: Patturajan, Meera,
 ; APPLICANT: Rothenberg, Mark,
 ; APPLICANT: Sciore, Paul,
 ; APPLICANT: Shenoy, Suresh,
 ; APPLICANT: Shimkets, Richard,
 ; APPLICANT: Si, Jingsheng,
 ; APPLICANT: Smithson, Glennda,
 ; APPLICANT: Spytek, Kimberly,
 ; APPLICANT: Stone, David,
 ; APPLICANT: Taupier, Raymond, jr.,
 ; APPLICANT: Tchernev, Velizar,
 ; APPLICANT: Vernet, Corine,
 ; APPLICANT: Zerhusen, Brian

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF

; FILE REFERENCE: 21402-222CIP

; CURRENT APPLICATION NUMBER: US/10/107,782

; PRIOR FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: 10/028,248

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/256,619

; PRIOR FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: 60/262,959

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/272,408

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 60/285,189

; PRIOR FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: 60/308,039

; PRIOR FILING DATE: 2001-07-26

; PRIOR APPLICATION NUMBER: 60/311,266

```
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 211
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-211

Query Match          94.3%; Score 1769; DB 15; Length 897;
Best Local Similarity 94.3%; Pred. No. 2.1e-183;
Matches 333; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 489 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 548
QY 61 RSPLOQYKLTDFKAREACANEATMATYNQLSYQKAKYHLCAGWLETGRVAYPTAFAS 120
Db 549 RSPLOQYKLTDFKAREACANEATMATYNQLSYQKAKYHLCAGWLETGRVAYPTAFAS 608
QY 121 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTXKVYVGDGFSYSGNLLQVLMSP 180
Db 609 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTXKVYVGDGFSYSGNLLQVLMSP 568
QY 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
Db 669 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 728
QY 241 VSMFFYNDLVNGTTLQTRVGSKLLITDRODPLHPTETRCVDRGTLEWDICASNGITHVI 300
Db 729 VSMFFYNDLVNGTTLQTRVGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIITHVI 788
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIIVTGAVALAAAYSIFRINRKTIGFXHF 353
Db 789 SRPLKAPPAPVTLTHTGLGAGIFFAILIIVTGAVALAAAYSIFRINRRTIGFQHF 841

RESULT 10
US-10-028-248A-2
; Sequence 2, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: NO. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
```

```
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2675
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (55)..(56)
; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (427)..(428)
; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid
US-10-028-248A-2

Query Match          83.3%; Score 1554; DB 15; Length 2675;
Best Local Similarity 83.9%; Pred. No. 3.1e-159;
Matches 296; Conservative 5; Mismatches 16; Indels 36; Gaps 1;

QY 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 2303 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 2362
QY 61 RSPLOQYKLTDFKAREACANEATMATYNQLSYQKAKYHLCAGWLETGRVAYPTAFAS 120
Db 2363 RSPLOQYKLTDFKAREACANEATMATYNQLSYQKAKYHLCAGWLETGRVAYPTAFAS 2422
QY 121 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMKDVNCTXKVYVGDGFSYSGNLLQVLMSP 180
Db 2423 QNCGSGVVGIVDYGPRPNKSEMDVFCYRMK----- 2453
QY 181 SLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
Db 2454 -----EVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHHLAN 2506
QY 241 VSMFFYNDLVNGTTLQTRVGSKLLITDRODPLHPTETRCVDRGTLEWDICASNGITHVI 300
Db 2507 VSMFFYNDLVNGTTLQTRVGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIITHVI 2566
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIIVTGAVALAAAYSIFRINRKTIGFXHF 353
Db 2567 SRPLKAPPAPVTLTHTGLGAGIFFAILIIVTGAVALAAAYSIFRINRRTIGFQHF 2619

RESULT 11
US-10-107-782-2
; Sequence 2, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
```



```
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shencoy, Suresh,
; APPLICANT: Shinkete, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zernusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 2675
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55)..(55)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (56)..(56)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (427)..(427)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-107-782-2
Query Match 83.3%; Score 1554; DB 15; Length 2675;
Best Local Similarity 83.9%; Pred. No. 3.1e-159;
Matches 296; Conservative 5; Mismatches 16; Indels 36; Gaps 1;
QY 1 MTGPGKHCKESHVYVDGNGINCEPEQLPDRCLQDNGQCHADAKCVDLHFQDQTTGVFHL 60
DB 2303 MTGPGKHCKESHVYVDGNGINCEPEQLPDRCLQDNGQCHADAKCVDLHFQDQTTGVFHL 2362
QY 61 RSPGQYKLTDFKAREACANEATWATYNQLSYXQKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 2363 RSPGQYKLTDFKAREACANEATWATYNQLSYXQKAKYHLCSAGWLETGRVAYPTAFAS 2422
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGYYVDGDFSYSGNLLQVIMSFP 180
DB 2423 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMK----- 2453
```

```
QY 181 SLTNFELTEVLAYSNSARGRAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
DB 2454 -----EVLAYSNSARGRAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 2506
QY 241 VSMFFYNLVNGTTLTQTLGSKLLITDRQDPLHPTETRCVDRDLEWDICASNGITHVI 300
DB 2507 VSMFFYNLVNGTTLTQTLGSKLLITASQDPLQPVQSRFVDGRAILOWDIFASNGIIRVI 2566
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIIVTGAVALAAAYSFRINRKTIGFXHP 353
DB 2567 SEPLKAPPAPVTLXHTGLGAGIFFCIIIVTGAVALAAAYSFRINRRTIGYQHF 2619
RESULT 12
US-09-842-930A-2
; Sequence 2, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-842-930A-2
Query Match 81.2%; Score 1514.5; DB 9; Length 1431;
Best Local Similarity 79.3%; Pred. No. 2.5e-155;
Matches 280; Conservative 29; Mismatches 43; Indels 1; Gaps 1;
QY 1 MTGPGKHCKESHVYVDGNGINCEPEQLPDRCLQDNGQCHADAKCVDLHFQDQTTGVFHL 60
DB 1008 MTGPGKHCKESHVYVDGNGINCEPEQLPDRCLQDNGQCHADAKCVDLHFQDQTTGVFHL 1067
QY 61 RSPGQYKLTDFKAREACANEATWATYNQLSYXQKAKYHLCSAGWLETGRVAYPTAFAS 120
DB 1068 RSPGQYKLTDFKAREACANEATWATYNQLSYXQKAKYHLCSAGWLETGRVAYPTAFAS 1127
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGYYVDGDFSYSGNLLQVIMSFP 180
DB 1128 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXKVGYYVDGDFSYSGNLLQVIMSFP 1187
QY 181 SLTNFELTEVLAYSNSARGRAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 240
DB 1188 SLTNFELTEVLAYSNSARGRAFLEHLTDLISIRGTLFVPQNSGLGENETLSGRDIEHHLAN 1247
QY 241 VSMFFYNLVNGTTLTQTLGSKLLITDRQDPLHPTETRCVDRDLEWDICASNGITHVI 300
DB 1248 VSMFFYNLVNGTTLTQTLGSKLLITASQDPLQPVQSRFVDGRAILOWDIFASNGIIRVI 1306
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIIIVTGAVALAAAYSFRINRKTIGFXHP 353
DB 1307 SEPLKAPPAPVTLXHTGLGAGIFFCIIIVTGAVALAAAYSFRINRRTIGYQHF 1359
RESULT 13
US-10-133-172-2
; Sequence 2, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
```

```
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-133-172-2

Query Match      81.2%; Score 1514.5; DB 14; Length 1431;
Best Local Similarity 79.3%; Pred. No. 2.5e-155;
Matches 280; Conservative 29; Mismatches 43; Indels 1; Gaps 1;

QY 1 MTGPGKHKECKSHYVGDGNGCPEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 1008 MTGPGKHKECKSHYVGDGVDCEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 1067

QY 61 RSLPGQYKLTDFKAREACANEATMATYNQLSYXOKAKYHLCAGWLETGRVAYPTAFAS 120
Db 1068 RSLPGQYKLTDFKAREACANEATMATYNQLSYXOKAKYHLCAGWLETGRVAYPTAFAS 1127

QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXXKVGYYGDFGFSYSGNLLQVLMSPF 180
Db 1128 QKCGANVGVIVDYGSRANKSEMDVFCYRMKDVNCTXXKVGYYGDFGFSYSGNLLQVLMSPF 1187

QY 181 SLTNFTLTVLAYSNSARGRAFLHLTDLISIRGTLFVFNQSGLGENETLSGRDIEHHLAN 240
Db 1188 SLTNFTLTVLAYSNSARGRAFLHLTDLISIRGTLFVFNQSGLGENETLSGRDIEHHLAN 1247

QY 241 VSMFFYNDLVNGTTLQTRIGSKLLITDRQDPLHPTETRCVDCGRDLEMDICASNGITHVI 300
Db 1248 VNSFFYNDLVNGTFLKIMLSGQLLITFSQQLH-QETRFVDGSRILQWDIIAANGLIHVI 1306

QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353
Db 1307 SBPLRAPPTAATAAHSLGLTGIFCAVVLVTGAVALAAYSFRILKQRTTGFQRF 1359

RESULT 14
US-10-028-248A-39
; Sequence 39, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
```

```
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods c
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 39
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-028-248A-39

Query Match      79.4%; Score 1480; DB 15; Length 1069;
Best Local Similarity 81.0%; Pred. No. 9.5e-152;
Matches 286; Conservative 4; Mismatches 27; Indels 36; Gaps 2;

QY 1 MTGPGKHKECKSHYVGDGNGCPEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 697 MTGPGKHKECKSHYVGDGNGCPEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 756

QY 61 RSLPGQYKLTDFKAREACANEATMATYNQLSYXOKAKYHLCAGWLETGRVAYPTAFAS 120
Db 757 RSLPGQYKLTDFKAREACANEATMATYNQLSYXOKAKYHLCAGWLETGRVAYPTAFAS 816

QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKDVNCTXXKVGYYGDFGFSYSGNLLQVLMSPF 180
Db 817 QNCGSGVGVIVDYGPRPNKSEMDVFCYRMKG-----SAGLFQQLSSRP 860

QY 181 SLTNFTLTVLAYSNSARGRAFLHLTDLISIRGTLFVFNQSGLGENETLSGRDIEHHLAN 240
Db 861 CIS-----RTPDDLIRGTLFVFNQSGLGENETLSGRDIEHHLAN 900

QY 241 VSMFFYNDLVNGTTLQTRIGSKLLITDRQDPLHPTETRCVDCGRDLEMDICASNGITHVI 300
Db 901 VSMFFYNDLVNGTTLQTRIGSKLLITASQDPLQPTETRFVDGRLQWDIFASNGLIHVI 960

QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353
Db 961 SRPLKAPPAPVTLTHTGLGAGIFFAIIILVTGAVALAAYSFRINRRTTIGFQHF 1013

RESULT 15
US-10-107-782-39
; Sequence 39, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
```

APPLICANT: Shenoy, Suresh.
APPLICANT: Shinkets, Richard,
APPLICANT: Si, Jingsheng,
APPLICANT: Smithson, Glennda,
APPLICANT: Spytek, Kimberly,
APPLICANT: Stone, David,
APPLICANT: Taupier, Raymond, jr.,
APPLICANT: Tchernev, Velizar,
APPLICANT: Vernet, Corine,
APPLICANT: Zerhusen, Brian
FILE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
FILE REFERENCE: 21402-222CIP
CURRENT APPLICATION NUMBER: US/10/107,782
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: 10/028,248
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/256,619
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/262,959
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/272,408
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/285,189
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: 60/308,039
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: 60/311,266
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/279,344
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 215
SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 39
LENGTH: 1069
TYPE: PRT
ORGANISM: Homo sapiens
US-10-107-782-39

Query Match 79.4%; Score 1480; DB 15; Length 1069;
Best Local Similarity 81.0%; Pred. No. 9.5e-152;
Matches 286; Conservative 4; Mismatches 27; Indels 36; Gaps 2;

QY 1 MTGPGKHCKESHVYGDGNGCEPEQLPDRCLQDNGQCHADAKCVDLHFQDITTVGVPHL 60
Db 697 MTGPGKHCKESHVYGDGNGCEPEQLPDRCLQDNGQCHADAKCVDLHFQDITTVGVPHL 756
QY 61 RSPLGQYKLTDFKAREACANEATMATYNQLSYQKAKYHLCAGWLETGRVAYPTAFAS 120
Db 757 RSPLGQYKLTDFKAREACANEATMATYNQLSYQKAKYHLCAGWLETGRVAYPTAFAS 816
QY 121 QNCGSGVGVIVDYGPRPNKSEMDVFCYMKDVNCTYKVYVGDGFSYGNLLQVLMSP 180
Db 817 QNCGSGVGVIVDYGPRPNKSEMDVFCYMKG-----SAGLFQQLSSRP 860
QY 181 SLTNFLTEVLAYSNSARGAFLEHLTDLSIRGTLFVPONSGIGENETLSGRDIEHHLAN 240
Db 861 CIS-----RTPDDLSIRGTLFVPONSGIGENETLSGRDIEHHLAN 900
QY 241 VSMFFYNDLVNGTTLQTRIGSKLLITDRQDPLHPTETRCVDRDITLWDICASNGITHVI 300
Db 901 VSMFFYNDLVNGTTLQTRIGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIHHVI 960
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXKIIIVTGAVALAAYSFRINRKTIGFXHF 353
Db 961 SRPLKAPPAPVTLTHTGLGAGIFPAILVTGAVALAAYSFRINRRTIGFQHF 1013

Search completed: March 30, 2004, 15:48:18
Job time : 46 secs

No art

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 30, 2004, 15:38:31 ; Search time 21 Seconds
(without alignments)
1616.934 Million cell updates/sec

Title: US-09-466-778b-11
Perfect score: 1865
Sequence: 1 MTGPGKHCKECSHYVDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1480	79.4	1069	T42681	hypothetical prote
2	264	14.2	275	JC6506	tumor necrosis fac
3	264	14.2	276	A47290	TSG-6 homolog PS4
4	262	14.0	277	A41735	hyaluronate-bindin
5	175	9.4	2109	I50421	aggreacan precursor
6	171	9.2	2124	A28452	proteoglycan core
7	170	9.1	2327	T42630	aggreacan - bovine
8	169.5	9.1	408	LKRT2	proteoglycan link
9	166.5	8.9	354	S42938	proteoglycan link
10	166	8.9	340	JC7505	brain link protein
11	165.5	8.9	355	LKCH	proteoglycan link
12	165	8.8	2132	A55182	aggreacan precursor
13	163.5	8.8	354	LKHU	proteoglycan link
14	162	8.7	2415	A39086	aggreacan precursor
15	161.5	8.7	354	S04243	proteoglycan link
16	160	8.6	370	S29139	aggreacan - pig (fr
17	158.5	8.5	1268	S52781	neurocan - mouse
18	157.5	8.4	371	A53908	brevican precursor
19	157.5	8.4	883	S49126	brevican precursor
20	157.5	8.4	1257	S28764	neurocan precursor
21	155.5	8.3	912	A54423	brevican precursor
22	154.5	8.3	862	S43922	versican - pig-tai
23	154.5	8.3	3562	A47171	chondroitin sulfat
24	151.5	8.1	1643	T14274	versican precursor
25	151.5	8.1	3381	T42389	versican precursor
26	150.5	8.1	883	S57653	brevican precursor
27	150.5	8.1	2409	A60979	brevican precursor
28	147	7.9	378	I46268	versican precursor
29	147	7.9	2397	A55535	versican precursor

30	139.5	7.5	1340	2	A39808	proteoglycan core
31	122	6.5	113	2	A55885	chondroitin sulfat
32	109.5	5.9	362	2	A35616	T-cell surface gly
33	109	5.8	1584	2	T22674	hypothetical prote
34	108.5	5.8	363	2	A37009	CD44 homolog membr
35	106	5.7	395	2	I77371	CD44R5 - human
36	104	5.6	359	2	S24240	lymphocyte surface
37	104	5.6	3507	2	T34513	hypothetical prote
38	102.5	5.5	365	2	A34424	CD44 membrane gly
39	101.5	5.4	1469	2	A36665	slit protein 2 pre
40	101.5	5.4	1480	2	A36665	slit protein 1 pre
41	100	5.4	2907	2	A57278	fibrillin-2 precu
42	99.5	5.3	728	2	I50719	C-Delta-1 - chicke
43	99	5.3	601	2	T22025	hypothetical prote
44	99	5.3	601	2	D89711	protein F40E10.4 [
45	98.5	5.3	3623	2	T09456	intrinsic factor-B

ALIGNMENTS

RESULT 1

T42681
hypothetical protein DKFp434E0321.1 - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T42681
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z22233
A:Accession: T42681
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-870; 871-1069 <AAA>
A:Cross-references: EMBL:AL133021
A:Experimental source: adult testis; clone DKFp434E0321
A>Note: the cDNA sequence contains a +1 frameshift near codon 870
C:Genetics:
A>Note: DKFp434E0321.1

Query Match 79.4%; Score 1480; DB 2; Length 1069;

Best Local Similarity 81.0%; Pred. No. 1.4e-117;

Matches 286; Conservative 4; Mismatches 27; Indels 36; Gaps 2;

QY	1	MTGPGKHCKECSHYVDGLNCEPQLPDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL	60
DB	697	MTGPGKHCKECSHYVDGLNCEPQLPDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL	756
QY	61	RSPLGQYKLTDFDKAREACANEATWATYNQLSYXKAKYHLCAGWLETGRVAYPTAFAS	120
DB	757	RSPLGQYKLTDFDKAREACANEATWATYNQLSYXKAKYHLCAGWLETGRVAYPTAFAS	816
QY	121	QNCGSVGVIVDYGPRPNKSEMMWVFCYRMKDVNCTXKVYVDGDFSYGNLLQVLMSPF	180
DB	817	QNCGSVGVIVDYGPRPNKSEMMWVFCYRMKG	860
QY	181	SUTNFLTVEVLAIVNSAGRAFLHLDLSIRGTLFVPQNSGLGENETLSGRDIEHLAN	240
DB	861	CIS	900
QY	241	VSMFFVNDLVNGTTIQLRGLSKLLITDRODPLHPPTETRCVDGRDITLWDICASNGTHVI	300
DB	901	VSMFFVNDLVNGTTIQLRGLSKLLITASQDPLQPTETRFVDGRALQWDIFASNGLIHVI	960
QY	301	SRXLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSIFRINRKTIGFXHF	353
DB	961	SRLKAPPAPVTLXHTGLGXGIFXIIILVTGAVALAAYSIFRINRKTIGFXHF	1013

RESULT 2

JC6506
tumor necrosis factor stimulated gene-6 protein - mouse
C:Species: Mus musculus (house mouse)

R.Lee, T.H.; Wisniewski, H.G.; Vilcek, J.
J. Cell Biol. 116, 543-557, 1992
A>Title: A novel secretory tumor necrosis factor-inducible protein (TSG-6) is a member of the interleukin-1 receptor type I family
A:Reference number: A41735; MUID:92112993; PMID:1730767
A:Accession: A41735
A:Molecule type: mRNA
A:Residues: 1-277 <LEX>
A:Cross-references: GB:M1165; NID:g339994; PIDN:AAB00792.1.; PID:g1332377
A:Experimental source: FS-4 fibroblast
A>Note: sequence extracted from NCBI backbone (NCBIN:76828, NCBIP:76833)
R:Wisniewski, H.G.; Burgess, W.H.; Oppenheim, J.D.; Vilcek, J.
Biochemistry 33, 7423-7429, 1994
A>Title: TSG-6, an arthritis-associated hyaluronan binding protein, forms a stable complex with hyaluronate
A:Reference number: A53642; MUID:94271799; PMID:7516184
A:Accession: D53642
A>Status: preliminary
A:Molecule type: protein
A:Residues: 18-27 <WIS>
C:Superfamily: C1r/C1s repeat homology; link protein repeat homology
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-277/Product: hyaluronate-binding protein TSG-6 #status predicted <MAT>
F:53-128/Domain: link protein repeat homology <LNK>
F:135-244/Domain: C1r/C1s repeat homology <CMR2>
F:118,258/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 14.0%; Score 262; DB 2; Length 277;
Best Local Similarity 43.3%; Pred No. 1e-14;
Matches 45; Conservative 16; Mismatches 43; Indels 0; Gaps 0;

QY 52 DTTVGVFHRLSPGLGYKLTDFDKAREACANEAAATMAATYNQLSYXQAKYHLCSAGMLETGR 111
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
32 ERAAGVYTHREARSGKYKLTYAEAKAVCEFGGHLATYKLEAARKIGFHVCAGGMWAKGR 91

QY 112 VAYPTAFASONGCGSWGIVDVGPRPNSEMDWVECYRMKVNC 155
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
92 VGPIVPKPGNGFGKTGIIDYGILNRSDWDAYCYNPHAKEC 135

RESULT 5
I50421
N;Alternate names: cartilage chondroitin sulfate proteoglycan core protein
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 21-Jan-2000
C;Accession: I50421; S39796; S27356; A25442; A32002; I50216; A37072; B37072
R:Li, H.; Schwartz, N.B.; Vertel, B.M.
J. Biol. Chem. 268, 23504-23511, 1993
A>Title: cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and aggrecan precursor - chicken
A:Reference number: A48884; MUID:94043149; PMID:8226678
A:Accession: I50421
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-2109 <LIX>
A:Cross-references: GB:L21913; NID:g416133; PIDN:AAB19128.1.; PID:g416134
R;Chandrasekaran, L.; Tanzer, M.L.
Biochem. J. 296, 885-887, 1993
A:Reference number: S39796; MUID:94107258; PMID:8280087
A:Contents: annotation; erratum
A:Accession: S39796
A:Molecule type: mRNA
A:Residues: 1-361, DL', 364-600, 'R', 602-999, 'R', 1001-1028, 'P', 1030-1250, 'D', 1252-1602, 'A', 'A'
A:Cross-references: GB:M88101
R;Chandrasekaran, L.; Tanzer, M.L.
Biochem. J. 288, 903-910, 1992
A>Title: Molecular cloning of chicken aggrecan. Structural analyses.
A:Reference number: S27356; MUID:93111968; PMID:1339285
A:Accession: S27356
A:Molecule type: mRNA
A:Residues: 1-361, 'DL', 364-600, 'R', 602-999, 'R', 1001-1028, 'P', 1030-1250, 'D', 1252-1549, 'T',
A:Cross-references: EMBL:M88101
R;Sai, S.; Tanaka, T.; Koshier, R.A.; Tanzer, M.L.
Proc. Natl. Acad. Sci. U.S.A. 83, 5081-5085, 1986

Query Match	9.4%;	Score 175;	DB 1;	Length 2109;
Best Local Similarity	23.4%;	Pred. No. 3.2e-06;		
Matches	75;	Conservative 31;	Mismatches 116;	Indels 98; Gaps 12;
QY	10	ECKSHYVGDCINCEPEOLPIDRCLDNGQCHA-----DAKCYVDLHFQDITTVG	56	
DB	194	QCDAGMLADQTVRYPIHLPERCYCGDKDFPGVRYTGVRETDYDYCYAEQMQCK	250	
QY	57	VFLHRSPLGQVKLTFFDKAREACANEATMATYNQLSYXQKAKYHLCAGWLETGRVAYPT	116	
DB	251	VFYATSP---EKFTFQBAFKCHSLGARLATTGELYLAWGDMDCSAGMLADRSVRYPI	307	
QY	117	AFASQNCGSGWGI--VDYQPR-----PNKSEMDWDFCYRMKDVNC-----TXKVG-VV	162	
DB	308	SRARPNCGGNLGVRYTYLVPANQTYGPHPSRDAICYSDGDFEALVPGLFIDEVGTEL	367	
QY	163	GDGFSY-----SNNLQVLMSPFSLNPL	186	
DB	368	GSAPTICTVTQTEVELPLPRNVNTEEBEARGSIATLEPMEITATATATELYEAFVLPDL--FA	425	
QY	187	TEVLAYSNSARGAFLEHLTDL-----SIRGTLFVPQNSGLGE-----	225	
DB	426	TSVTVEFASPREENVNTEETIGIMAVPEVTVTSVSGTAF---TTGMAEVSVEEIAVTA	482	

RESULT 7

T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence revision 11-Jan-2000 #text change 04-Mar-2000

```

5 GKHKCECKSHVYVGDGLNCEPEQLPIDRCLQDNQGHADAKVDLHFQDTTGVGFHLRSPL 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
189 GRKYCE : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 GQYKLTDFDKAREACANAATWATYNQLSYQKAKYHLCGAGWLETGRVAYPTAFASQNCG 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
222 GRYNLNFHEARQAQLDQDAVIAFSDQLYDAWRGGLDWCNAGWLSGDSGVQYPTITKREPFCG 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 --SGVVGIVDYGPRNKSEMDVFCY 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
282 GQNTVPGVRNRYGFWDKSKRYDVFCE 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
S42938
Proteoglycan link protein precursor - horse
N:Alternate names: cartilage link protein
C:Species: Equus caballus (domestic horse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Jun-1999
C:Accession: S42938
R:Dudhia, J.; Platt, D.
submitted to the EMBL Data Library, March 1994
A:Reference number: S42938
A:Accession: S42938
A:Molecule type: mRNA
A:Residues: 1-354 <DUD>
A:Cross-references: EMBL:X78077; NID:G459438; PIDN:CAA54987.1; PID:9459439
C:Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat
C:Keywords: cartilage; duplication; extracellular matrix; glycoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-354/Product: proteoglycan link protein #status predicted <MAT>
F:54-141/Domain: immunoglobulin homology <IMM>
F:176-253/Domain: link protein repeat homology <LNK1>
F:274-350/Domain: link protein repeat homology <LNK2>
F:21-56/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:61-139/Disulfide bonds: #status predicted

Query Match 8.9%; Score 166.5; DB 1; Length 354;
Best Local Similarity 30.1%; Pred. NO. 1.8e-06;
Matches 44; Conservative 22; Mismatches 51; Indels 29; Gaps 4;

QY 5 GKHKCECKSHVYVGDGLNCEPEQLPIDRCLQDNQGHADAKVDLHFQDTTGVGFHLRSPL 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
135 GRKYCE : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 GQYKLTDFDKAREACANAATWATYNQLSYQKAKYHLCGAGWLETGRVAYPTAFASQNCG 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
168 GRYNLNFHEARQAQLDQDAVIAFSDQLYDAWRGGLDWCNAGWLSGDSGVQYPTITKREPFCG 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 --SGVVGIVDYGPRNKSEMDVFCY 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
228 GQNTVPGVRNRYGFWDKSKRYDVFCE 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
JC7505
brain link protein-1 - human
C:Species: Homo sapiens (man)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C:Accession: JC7505
R:Hirakawa, S.; Ohashi, T.; Su, W.D.; Yoshioka, H.; Murakami, T.; Arata, J.; Ninomiya,
Biochem. Biophys. Res. Commun. 276, 982-989, 2000
A:Title: The brain link protein-1 (BRAL1): cDNA cloning, genomic structure, and character:
A:Reference number: JC7505
A:Accession: JC7505
A:Molecule type: mRNA
A:Residues: 1-340 <HIR>
A:Cross-references: DDBJ:AB049054
C:Comment: This protein stabilizes the binding between hyaluronan and brain-specific hya
C:Genetics:
A:Gene: bral1
C:Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat
C:Keywords: brain; tandem repeat

```

```

Query Match      8.9%; Score 166; DB 2; Length 340;
Best Local Similarity 31.7%; Pred. No. 1.9e-06;
Matches 40; Conservative 21; Mismatches 51; Gaps 4;

QY 33 LODNQCHADAKCVDLH-FQDTTVG-----VFHLRSPLGQYKLTDPKAREACANERA 83
DB 120 LDEGR-----YRCEILINGEDSVALTSLGQVFPYQPSRGYQFNYYEAKACEBQDG 175
QY 84 TWATYNQLSYXQKAKYHLCAGWLETGRVAYPTAFASQCGS-GVVGIVDYDGPKNKSM 142
DB 176 RLATYSQLYQAMTEGLDWCNAGWLLGVSRYVPVLTARPCGGRGPRGIRSYGPRDMRDR 235
QY 143 WDVFCY 148
DB 236 YDAFCF 241

proteoglycan link protein precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text_change 22-Jun-1999
C:Accession: A28305; A24881; A39097; B39097
R:Kiss, J.; Deak, F.; Westric, S.; Bellius, H.; Soos, J.; Dekany, K.; Argraves, W.S.; Spa
Proc. Natl. Acad. Sci. U.S.A. 84, 6399-6403, 1987
A:Title: Structure of the chicken link protein gene: exons correlate with the protein do
A:Reference number: A28305; MUID:87317659; PMID:3476955
A:Accession: A28305
A:Molecule type: DNA
A:Residues: 1-355 <KIS>
A:Cross-references: GB:M35038; NID:g212264; PIDN:AAA48941.1; PID:g212267
R:Deak, F.; Kiss, J.; Sparks, K.J.; Argraves, W.S.; Hampikian, G.; Goetinck, P.F.
Proc. Natl. Acad. Sci. U.S.A. 83, 3766-3770, 1986
A:Title: Complete amino acid sequence of chicken cartilage link protein deduced from cDN
A:Reference number: A24881; MUID:86233315; PMID:3459154
A:Accession: A24881
A:Molecule type: mRNA
A:Residues: 1-355 <DEA>
A:Cross-references: GB:M13212; NID:g212259; PIDN:AAA48940.1; PID:g212260
A:Experimental source: embryonic sternal cartilage
R:Wu, L.N.Y.; Genge, B.R.; Wuthier, R.E.
J. Biol. Chem. 266, 1187-1194, 1991
A:Title: Association between proteoglycans and matrix vesicles in the extracellular matr
A:Reference number: A39097; MUID:91033230; PMID:1985942
A:Accession: A39097
A:Molecule type: protein
A:Residues: 40-55, 'X', 57-60, 'X', 62-75, 'X', 77-78, 'X' <WUA>
A:Note: 38K protein, a major component of matrix vesicles; Asn-56 appears to be glycosyl
A:Accession: B39097
A:Molecule type: protein
A:Residues: 40-55, 'X', 57-59 <WU2>
A:Note: 39K protein, a minor component of matrix vesicles
C:Comment: Link proteins stabilize the aggregates of proteoglycan monomers with hyaluron
structures that induce mineralization).
C:Genetics:
A:Introns: 34/1; 159/1; 260/1
C:Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat h
C:Keywords: cartilage; duplication; extracellular matrix; glycoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-385/Product: link protein #status predicted <WAT>
F:54-142/Domain: immunoglobulin homology <IMW>
F:177-254/Domain: link protein repeat homology <LNK2>
F:275-351/Domain: link protein repeat homology <LNK2>
F:21/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:56/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:61-140,182-253,206-227,280-350,305-326/Diulfide bonds: #status predicted

Query Match      8.9%; Score 165.5; DB 1; Length 355;
Best Local Similarity 29.9%; Pred. No. 2.2e-06;
Matches 44; Conservative 26; Mismatches 46; Indels 31; Gaps 5;

```

```

QY 5 GKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHLRSP- 63
DB 136 GRYKOE-----VIEGL-----EDTAVVALLNLE-----GVVFPYSPR 167
QY 64 LGQYKLTEDKAREACANEAATWATYNQLSYXQKAKYHLCAGWLETGRVAYPTAFASQNC 123
DB 168 LGRYNLNFEHAQACLDDQDSIIASFQDLYEARSLGDLWCNAGWLDGSGVQYPIITPREPC 227
QY 124 G--SGVGVIVDYDGPKNKSEMMDVFCY 148
DB 228 GGNTPVGRVNYGFWDKERSRYDVFCF 254

RESULT 12
A55182
aggrecan precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 21-Jan-2000
C:Accession: A55182; S55329; S50207; S51355; I78532; I58123
R:Walcz, E.; Deak, F.; Erhardt, P.; Coulter, S.N.; Fuloep, C.; Horvath, P.; Doege, K.J.
Genomics 22, 364-371, 1994
A:Title: Complete coding sequence, deduced primary structure, chromosomal localization,
A:Reference number: A55182; MUID:95104847; PMID:7806222
A:Accession: A55182
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-2132 <WAL>
A:Cross-references: GB:I07049; NID:g678541; PIDN:AAC37670.1; PID:g191772
R:Watanabe, H.; Gao, L.; Sugiyama, S.; Doege, K.; Kimata, K.; Yamada, Y.
Biochem. J. 308, 433-440, 1995
A:Title: Mouse aggrecan, a large cartilage proteoglycan: protein sequence, gene structur
A:Reference number: S55329; MUID:95289972; PMID:7772024
A:Accession: S55329
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-567, 'G', 569-1728, 'I', 1730-2132 <WAT1>
A:Cross-references: GB:U22901; NID:9886014
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
A:Reference number: S50206; MUID:95035091; PMID:7524681
A:Accession: S50207
A:Molecule type: mRNA
A:Residues: 350-481, 'R', 483-506 <GLU1>
A:Cross-references: EMBL:X80279; NID:9673432
R:Glumoff, V.
submitted to the EMBL Data Library, July 1994
A:Reference number: S51355
A:Accession: S51355
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 350-383, 'CPVMSQRRPMAA' <GLU2>
A:Cross-references: EMBL:X80279
R:Watanabe, H.; Kimata, K.; Line, S.; Strong, D.; Gao, L.Y.; Kozak, C.A.; Yamada, Y.
Nature Genet. 7, 154-157, 1994
A:Title: Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the aggre
A:Reference number: I58123; MUID:95004579; PMID:7920633
A:Accession: I78532
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 211-326 <WAT2>
A:Cross-references: GB:S73722; NID:g765215; PIDN:AAB32160.1; PID:g765216
A:Accession: I58123
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 211-240, 'MCTASLRWRVRSFMRHPQNRSPRQPTS', 'AGGCHAWPPQASSTWPGRAVWTCALAGW', 'R
A:Cross-references: GB:S73720; NID:g765211; PIDN:AAB32159.1; PID:g765212
C:Genetics:
A:Map position: 7
A:Introns: 253/1
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C:Keywords: cartilage; extracellular matrix
F:1-19/Domain: signal sequence #status predicted <SIG>

```


RESULT 14

A39086

aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan; p1
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #Sequence_revision 01-Dec-2000 #text change 08-Dec-2000
C;Accession: A39086; S50206; A43919; S46659; S66389; S68646; S62786; A34226; B43919; C433
R;Doerge, K.J.; Sasaki, M.; Kimura, T.; Yamada, Y.
J. Biol. Chem. 266, 894-902, 1991
A;Title: Complete coding sequence and deduced primary structure of the human cartilage 1
A;Reference number: A39086; MUID:91093289; PMID:1985970
A;Accession: A39086
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2162, 2201-2329, 'A', 2392-2415 <DOB>
A;Cross-references: GB:M55172; NID:G178258; PIDN:AAAG62824.1; PID:G178259
E;Glumoff, V.; Savontaus, M.; Venhanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by R
A;Reference number: S50206; MUID:95035091; PMID:7524681
A;Accession: S50206
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 350-497 <GLU>
A;Cross-references: EMBL:X80278; NID:G516295
A;Note: this translation is not annotated in GenBank entry HSAGGREC, release 113.0
E;Sandy, J.D.; Flannery, C.R.; Neame, P.J.; Lohmander, L.S.
J. Clin. Invest. 89, 1512-1516, 1992
A;Title: The structure of aggrecan fragments in human synovial fluid. Evidence for the i
domain.
A;Reference number: A43919; MUID:92235266; PMID:1569188
A;Accession: A43919
A;Molecule type: protein
A;Residues: 361-370, 'X', 372-373; 393-399, 'X', 401-407, 'X', 409 <SAND>

A;Cross-references: PIDN:AA22079.1; PID:9248844; PIDN:AA22077.1; PID:9248842; PIDN:AAH
A;Experimental source: synovial fluid
A;Note: sequences modified after extraction from NCBI backbone
R;Barry, F.P.; Neame, P.J.; Sasse, J.; Pearson, D.
Matrix Biol. 14, 323-328, 1994
A;Title: Length variation in the keratan sulfate domain of mammalian aggrecan.
A;Reference number: 146998; MUID:95128522; PMID:7827755
A;Accession: S46659
A;Molecule type: DNA
A;Residues: 764-765, 'A', 767-846, 'V', 848-862, 'X', 864 <BAR>
A;Cross-references: EMBL:S74659; NID:9807127; PIDN:AA22077.1; PID:9807128
A;Note: the authors translated the codon GAA for residue 803 as Ala and Cyt for residue
F;Ilic, M.Z.; Mok, M.T.; Williamson, O.D.; Campbell, M.A.; Hughes, C.E.; Handley, C.J.
Arch. Biochem. Biophys. 322, 22-30, 1995
A;Title: Catabolism of aggrecan by explant cultures of human articular cartilage in the
A;Reference number: S66389; MUID:96004775; PMID:7574678
A;Accession: S66389
A;Status: preliminary
A;Molecule type: protein
A;Residues: 17-23;24, 'X', 26-27;393-401;402-403 <ILI>
R;Fosang, A.J.; Last, K.; Knauper, V.; Murphy, G.; Neame, P.J.
FEBS Lett. 380, 17-20, 1996
A;Title: Degradation of cartilage aggrecan by collagenase-3 (MMP-13).
A;Reference number: S68646; MUID:96181659; PMID:8603731
A;Accession: S68646
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'V', 404-405, 'XX' <FOS>
R;Dudhia, J.; Davidson, C.M.; Wells, T.M.; Vynios, D.H.; Hardingham, T.E.; Bayliss, M.T.
Biochem. J. 313, 933-940, 1996
A;Title: Age-related changes in the content of the C-terminal region of aggrecan in huma
A;Reference number: S62786; MUID:96190740; PMID:8611178
A;Accession: S62786
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1778-1927, 'A', 1929-1963, 'V', 1965-2162, 2201-2415 <DUD>
A;Cross-references: EMBL:X17406; NID:930248; PIDN:CAA35463.1; PID:930249
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1990
R;Baldwin, C.T.; Reginato, A.M.; Prockop, D.J.
J. Biol. Chem. 264, 15747-15750, 1989
A;Title: A new epidermal growth factor-like domain in the human core protein for the lat
A;Reference number: A34226; MUID:89380154; PMID:2789216
A;Accession: A34226
A;Molecule type: mRNA
A;Residues: 1936-1963, 'V', 1965-2069, 'A', 2071-2415 <BAL>
A;Cross-references: GB:J05062; NID:9181167; PIDN:AAA35726.1; PID:9181168
C;Genetics:
A;Gene: GDB:AGC3; CSPG1; CSPGCP; MSK16
A;Cross-references: GDB:127479; OMIM:155760
A;Map position: 15q26-15q26
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: alternative splicing; cartilage; chondroitin sulfate proteoglycan; extracell
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2415/Product: aggrecan cartilage long splice form #status predicted <MATL>
F;20-2162,2201-2415/Product: aggrecan cartilage short splice form #status predicted <MAT
F;20-2162,2201-2329, 'A', 2332-2415/Product: aggrecan short splice form #status predicted
F;44-135/Domain: immunoglobulin homology <IMM>
F;170-247/Domain: link protein repeat homology <LNK1>
F;268-349/Domain: link protein repeat homology <LNK2>
F;495-572/Domain: link protein repeat homology <LNK3>
F;593-673/Domain: link protein repeat homology <LNK4>
F;677-861/Domain: keratan sulfate attachment #status predicted <KSA>
F;864-1510/Domain: chondroitin sulfate attachment #status predicted <CS1>
F;1511-2162/Domain: chondroitin sulfate attachment #status predicted <CS2>
F;2168-2198/Domain: EGF homology <EGF>
F;2205-2325/Domain: C-type lectin homology <LCH>
F;2332-2388/Domain: complement factor H repeat homology <FHD>
F;126,239,333,387,434,602,657,737,1898/Binding site: carbohydrate (Asn) (covalent) #stat
F;371,376/Binding site: keratan sulfate (Thr) (covalent) #status predicted

Query Match 8.7%; Score 162; DB 1; Length 2415;
Best Local Similarity 26.5%; Pred. No. 4.9e-05;
Matches 63; Conservative 25; Mismatches 82; Indels 68; Gaps 10;

QY 47 DLHFQDFTVG-----VFHLRSLPGQYKLFDFKAREACANEAAATATYNQLSYXQKA 97
Db 460 DLVVQVTVAPGQPHLPQGVVPHRPGTRISLTFEEAQACPGCTGAVIASPEQLQAAVEA 519
QY 98 KYHLCSAGNLETGRVAYPTAFASQNC---SGVGVIVDYGRPNKSEMWDFVFCY----- 148
Db 520 GYEQCDAGWLRLDQTVRPIVSPRTPCVGDKSSSGVATYGVRRPS-TETDYDVCVDRLEG 578
QY 149 -----RMKDV-----NCTKXGVYVGDGFS-----YSGNLLQVIMSFPSLT- 183
Db 579 EVFFATRLQFTFOEALFECESHNATATTQGLYAAMSRGLDKCYAGLADGSLRPIVTP 638
QY 184 -----NFIETVLAYSNNS-----ARGRAFLHLTDLISIRGTLFVPQNSGLGENE 227
Db 639 RPACGDKPGVTRVYLPNQTGLPDLRHHAF-----CFRGISAVPSP---GEE 686

RESULT 15

S04243
proteoglycan link protein precursor - pig
N;Alternate names: cartilage link protein
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 22-Jun-1999
C;Accession: S04243; I47145
R;Dudhia, J.; Hardingham, T.E.
J. Mol. Biol. 206, 749-753, 1989
A;Title: Appendix. Isolation and sequence of cDNA clones for pig and human cartilage lin
A;Reference number: S04243
A;Accession: S04243
A;Molecule type: mRNA
A;Residues: 1-354 <DUD>
A;Cross-references: EMBL:Y00165; NID:92009; PIDN:CAA68358.1; PID:92010
R;Perkins, S.J.; Neallis, A.S.; Dudhia, J.; Hardingham, T.E.
J. Mol. Biol. 206, 737-753, 1989
A;Title: Immunoglobulin fold and tandem repeat structures in proteoglycan N-terminal dom
A;Reference number: I47145; MUID:89293837; PMID:2738916
A;Accession: I47145
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-354 <PER>
A;Cross-references: EMBL:Y00165; NID:92009; PIDN:CAA68358.1; PID:92010
C;Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat h
C;Keywords: cartilage; duplication; extracellular matrix; glycoprotein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-354/Product: proteoglycan link protein #status predicted <MAT>
F;54-141/Domain: immunoglobulin homology <IMM>
F;176-253/Domain: link protein repeat homology <LNK1>
F;2176-350/Domain: link protein repeat homology <LNK2>
F;21,56/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;61-139/Disulfide bonds: #status predicted

Query Match 8.7%; Score 161.5; DB 1; Length 354;
Best Local Similarity 29.5%; Pred. No. 4.8e-06;
Matches 43; Conservative 22; Mismatches 52; Indels 29; Gaps 4;

QY 5 GKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDFTVGVFHLRSLP 64
Db 135 GRYKCE-----VIEGL-----EDDTAVVALDEGV---VFYFPRL 167
QY 65 GOYKLTDFKAREACANEAAATATYNQLSYXQKAKYHLCAGMLETGRVAYPTAFASQNC 124
Db 168 GRYNLFHFAQAQCLDQDAVIAFDFOLYAWRGGLDWCNAGMLSDGSVQYPTTKPREPG 227
QY 125 --SGVGVIVDYGRPNKSEMWDFVFCY 148
Db 228 GQNTVPGVNRVGFWDKDKSRDYVFCF 253

Search completed: March 30, 2004, 15:43:02
Job time : 22 secs